



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN RE APPLICATION OF: Hiroshi TAIRA, et al.

GAU:

SERIAL NO: 10/645,623

EXAMINER:

FILED: August 22, 2003

FOR: PROGRAM, SYSTEM AND METHOD FOR ANALYZING RETRIEVAL KEYWORD

REQUEST FOR PRIORITY

COMMISSIONER FOR PATENTS
ALEXANDRIA, VIRGINIA 22313

SIR:

Full benefit of the filing date of U.S. Application Serial Number , filed , is claimed pursuant to the provisions of 35 U.S.C. §120.

Full benefit of the filing date(s) of U.S. Provisional Application(s) is claimed pursuant to the provisions of 35 U.S.C. §119(e): Application No. Date Filed

Applicants claim any right to priority from any earlier filed applications to which they may be entitled pursuant to the provisions of 35 U.S.C. §119, as noted below.

In the matter of the above-identified application for patent, notice is hereby given that the applicants claim as priority:

<u>COUNTRY</u>	<u>APPLICATION NUMBER</u>	<u>MONTH/DAY/YEAR</u>
JAPAN	2002-243466	August 23, 2002
JAPAN	2002-382396	December 27, 2002
JAPAN	2003-070281	March 14, 2003

Certified copies of the corresponding Convention Application(s)

are submitted herewith

will be submitted prior to payment of the Final Fee

were filed in prior application Serial No. filed

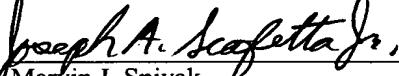
were submitted to the International Bureau in PCT Application Number
Receipt of the certified copies by the International Bureau in a timely manner under PCT Rule 17.1(a) has been acknowledged as evidenced by the attached PCT/IB/304.

(A) Application Serial No.(s) were filed in prior application Serial No. filed ; and

(B) Application Serial No.(s)

- are submitted herewith
- will be submitted prior to payment of the Final Fee

Respectfully Submitted,

OBLON, SPIVAK, McCLELLAND,
MAIER & NEUSTADT, P.C.


Marvin J. Spivak

Registration No. 24,913

Joseph A. Scafetta, Jr.
Registration No. 26,803

Customer Number

22850

Tel. (703) 413-3000
Fax. (703) 413-2220
(OSMMN 05/03)

10/645,623
0351021-1

13

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 Date of Application: 2002年 8月23日

出願番号 Application Number: 特願2002-243466

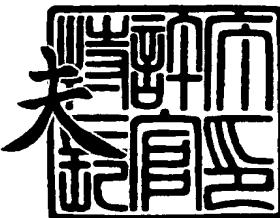
[ST. 10/C]: [JP2002-243466]

出願人 Applicant(s): 株式会社東芝

2003年 9月 2日

特許庁長官
Commissioner,
Japan Patent Office

今井康



出証番号 出証特2003-3071541

【書類名】 特許願
【整理番号】 A000203825
【提出日】 平成14年 8月23日
【あて先】 特許庁長官 殿
【国際特許分類】 G06F 17/00
【発明の名称】 検索キーワード分析プログラム及びシステム並びに方法
【請求項の数】 14
【発明者】
【住所又は居所】 東京都府中市東芝町1番地 株式会社東芝府中事業所内
【氏名】 平 博司
【発明者】
【住所又は居所】 東京都府中市東芝町1番地 株式会社東芝府中事業所内
【氏名】 平井 潤
【特許出願人】
【識別番号】 000003078
【氏名又は名称】 株式会社 東芝
【代理人】
【識別番号】 100058479
【弁理士】
【氏名又は名称】 鈴江 武彦
【電話番号】 03-3502-3181
【選任した代理人】
【識別番号】 100084618
【弁理士】
【氏名又は名称】 村松 貞男
【選任した代理人】
【識別番号】 100068814
【弁理士】
【氏名又は名称】 坪井 淳

【選任した代理人】

【識別番号】 100092196

【弁理士】

【氏名又は名称】 橋本 良郎

【選任した代理人】

【識別番号】 100091351

【弁理士】

【氏名又は名称】 河野 哲

【選任した代理人】

【識別番号】 100088683

【弁理士】

【氏名又は名称】 中村 誠

【選任した代理人】

【識別番号】 100070437

【弁理士】

【氏名又は名称】 河井 将次

【手数料の表示】

【予納台帳番号】 011567

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 検索キーワード分析プログラム及びシステム並びに方法

【特許請求の範囲】

【請求項 1】 コンピュータに、

WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、前記アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録するアクセスシーケンス抽出機能と、

前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けられている場合に、前記検索キーワードを抽出し、前記検索キーワードの抽出結果を示すキーワードデータを前記記録手段に記録するキーワード抽出機能と、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する質的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードのうち前記アクセスシーケンスにおいて入力された検索キーワードと前記質的評価値との対応関係を示す関係データを前記記録手段に記録する評価値算出機能と

を実現させるための検索キーワード分析プログラム。

【請求項 2】 請求項1記載の検索キーワード分析プログラムにおいて、

前記評価値算出機能は、前記質的評価値に加えて、前記検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する量的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードのうち前記アクセスシーケンスにおいて入力された検索キー

ワードと前記質的評価値と前記量的評価値の対応関係を示す関係データを前記記録手段に記録する

ことを特徴とする検索キーワード分析プログラム。

【請求項3】 コンピュータに、

WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、前記アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録するアクセスシーケンス抽出機能と、

前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けられており前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスにおいて前記検索キーワードが入力されている場合に、前記検索キーワードを抽出し、前記検索キーワードの抽出結果を示すキーワードデータを前記記録手段に記録するキーワード抽出機能と、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する質的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードと前記質的評価値との対応関係を示す関係データを前記記録手段に記録する評価値算出機能と

を実現させるための検索キーワード分析プログラム。

【請求項4】 請求項3記載の検索キーワード分析プログラムにおいて、

前記評価値算出機能は、前記質的評価値に加えて、前記検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する量的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードと前記質的評価値と前記量的評価値の対応関係を示す関係データ

ータを前記記録手段に記録する
ことを特徴とする検索キーワード分析プログラム。

【請求項 5】 請求項 2 又は請求項 4 記載の検索キーワード分析プログラムにおいて、

コンピュータに、前記記録手段に記録された関係データにしたがって、前記質的評価値と前記量的評価値に基づくグラフを描画するグラフ化機能を実現させるためのプログラムを附加したことを特徴とする検索キーワード分析プログラム。

【請求項 6】 請求項 2、請求項 4 又は請求項 5 記載の検索キーワード分析プログラムにおいて、

前記量的評価値は、前記検索キーワードの入力数が多いほど値が大きくなる検索結果のニーズ値であり、

前記量的評価規則は、前記検索キーワードの入力数が多いほど前記ニーズ値が増加する規則であることを特徴とする検索キーワード分析プログラム。

【請求項 7】 請求項 1 乃至請求項 6 のいずれか記載の検索キーワード分析プログラムにおいて、

前記Web アクセスログは、WWWにおけるアクセス毎に、前記アクセス要求側識別データと前記アクセス先識別データとアクセス時刻を示すデータとを関係付けており、

前記質的評価規則は、前記アクセスシーケンスにおける前記検索キーワード入力後のアクセスから最後のアクセスまでの時間間隔が長いほど前記アクセス要求側の満足度が増加する規則であり、

前記質的評価値は、前記満足度である
ことを特徴とする検索キーワード分析プログラム。

【請求項 8】 請求項 1 乃至請求項 6 のいずれか記載の検索キーワード分析プログラムにおいて、

前記質的評価規則は、分析候補の検索キーワードと前記分析候補の検索キーワードの入力後にアクセスされた場合に成功度を増加するアクセス先識別データとを関連付けたデータと前記記録手段に記録されたキーワードデータとにに基づいて

、前記アクセスシーケンスにおいて前記分析候補の検索キーワードが入力されておりその入力後前記分析候補の検索キーワードと関連付けされているアクセス先識別データの示すアクセス先がアクセスされている場合に、前記アクセスシーケンスに対する成功度を増加する規則であり、

前記質的評価値は、前記成功度であることを特徴とする検索キーワード分析プログラム。

【請求項9】 請求項8に記載の検索キーワード分析プログラムにおいて、前記質的評価規則は、前記分析候補の検索キーワードが入力されてから前記分析候補の検索キーワードと関連付けされているアクセス先識別データの示すアクセス先がアクセスされるまでのアクセス数が少ないほど加えられる成功度を大きくすることを特徴とする検索キーワード分析プログラム。

【請求項10】 請求項1乃至請求項9のいずれか記載の検索キーワード分析プログラムにおいて、

前記Webアクセスログは、前記アクセス先識別データの示すアクセス先がハイパーリンクを使ってアクセスされた場合に、前記アクセス先識別データの示すアクセス先の前段でアクセスされたアクセス先を示すレファラ情報と前記アクセス先識別データとを関連付けたデータを含み、

前記アクセスシーケンス抽出機能は、前記レファラ情報が前記検索システムを示す場合に、前記レファラ情報に関連付けされている前記アクセス先識別データの示すアクセス先へのアクセスを含むアクセスシーケンスを抽出し、

前記キーワード抽出機能は、前記レファラ情報が前記検索システムを示す場合に、前記レファラ情報に含まれている検索キーワードを抽出し、

前記評価値算出機能は、前記アクセスシーケンスに対する前記検索キーワード入力後の質的評価値を算出する

ことを特徴とする検索キーワード分析プログラム。

【請求項11】 WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出するアクセスシーケンス

抽出手段と、

前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けされている場合に、前記検索キーワードを抽出するキーワード抽出手段と、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記アクセスシーケンスに対応する質的評価値を算出し、前記アクセスシーケンスにおいて入力された検索キーワードと前記質的評価値との対応関係を示す関係データを作成する評価値算出手段と

を具備する検索キーワード分析システム。

【請求項12】 WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出するアクセスシーケンス抽出手段と、

前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けられており前記アクセスシーケンスにおいて前記検索キーワードが入力されている場合に、前記検索キーワードを抽出するキーワード抽出手段と、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記アクセスシーケンスに対応する質的評価値を算出し、前記検索キーワードと前記質的評価値との対応関係を示す関係データを作成する評価値算出手段と
を具備する検索キーワード分析システム。

【請求項13】 WWWにおける検索システムに入力された検索キーワードについて分析するためにコンピュータにより実施される検索キーワード分析方法において、

WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アセ

ス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、前記アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録するとともに、前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けされている場合に、前記検索キーワードを抽出し、前記検索キーワードの抽出結果を示すキーワードデータを前記記録手段に記録し、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する質的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードのうち前記アクセスシーケンスにおいて入力された検索キーワードと前記質的評価値との対応関係を示す関係データを前記記録手段に記録することを特徴とする検索キーワード分析方法。

【請求項14】 WWWにおける検索システムに入力された検索キーワードについて分析するためにコンピュータにより実施される検索キーワード分析方法において、

WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、前記アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録し、

前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けられており前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスにおいて前記検索キーワードが入力されている場合に、前記検索キーワードを抽出し、前記検索キーワードの抽出結果を示すキーワードデータを前記記録手段に記録し、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する質的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードと前記質的評価値との対応関係を示す関係データを前記記録手段に記録することを特徴とする検索キーワード分析方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、WWW (World Wide Web) において、目的のWebページ又はWebサイトを探すために、Web閲覧者（以下、単に「閲覧者」という）からWWW検索システム（以下、「検索システム」という）に入力された検索キーワードの特徴を分析するためのプログラム及びシステム並びに方法に関する。

【0002】

【従来の技術】

インターネットにおけるアクセス先の例として、Webページ又はWebサイトがある。

【0003】

インターネット上のWebページ又はWebサイトは、宣伝広告、情報発信、受注受付などの様々な目的のために運営される。かかる目的を効果的に達成するために、Webページ又はWebサイトが分析、調査される。

【0004】

Webページ又はWebサイトの分析、調査の方法として、どのWebページが、いつ、どのようにアクセスされたかを記録するWebアクセスログを用いたWebアクセスログ分析がある。

【0005】

Webアクセスログには、主に2種類ある。

【0006】

第1のWebアクセスログは、单一又は複数のWebサーバに記録されたWe

Webアクセスログ（狭義のWebアクセスログ）である。

【0007】

第2のWebアクセスログは、一人又は複数の閲覧者による様々なWebページ又はWebサイトへのアクセスを記録したWebアクセスログである。この第2のWebアクセスログには、例えば、WWW視聴率調査会社がパネル（モニター）から収集した、各パネルによるWebページ又はWebサイトへのアクセスログがある。

【0008】

以下において、Webアクセスログとは、第1のWebアクセスログを意味するとして説明する。

【0009】

例えば、Webアクセスログ分析では、Webページ又はWebサイトにアクセスした閲覧者数（ユーザ数）が、時間毎、曜日毎にどのように変化するか調査される。

【0010】

また、Webアクセスログ分析では、閲覧者が使用しているブラウザ又はコンピュータのOSの種類が調査される場合もある。

【0011】

また、Webアクセスログ分析では、一つのWebサイト内又は複数のWebサイト間をまたがって閲覧者がどのような閲覧行動をとるかについて調査される場合もある。例えば、閲覧者がWebページAを閲覧し、次にWebページBを閲覧したことを明らかにするなどのような経路分析もWebアクセスログ分析に含まれる。

【0012】

一方、インターネット上のWebページ又はWebサイトの閲覧を補助するために、閲覧者の入力した検索キーワード（例えば、「車」「ペット」など）に応じて、この検索キーワードに関する内容のWebページ（以下、「キーワード関連ページ」という）又はWebサイト（以下、「キーワード関連サイト」という）についての情報（例えば、キーワード関連ページ又はキーワード関連サイトの

U R L (Uniform Resource Locators)、要約情報、リンク情報など) のリストを検索結果ページとして表示する検索システムが存在する。この検索システムは、広く利用されている。

【0013】

図13は、検索システムを用いた代表的なWebページの遷移の例を示す。

【0014】

検索キーワード入力画面1上で閲覧者のブラウザから検索キーワードが入力され、検索開始が指示されると、検索システムは、キーワード関連ページへのリンク2a～2cを抽出し、抽出されたキーワード関連ページへのリンク2a～2cを表示する検索結果ページ2を閲覧者に対して表示する。

【0015】

検索結果ページ2上のいずれかのリンクがクリックされると、このクリックされたリンクの先のキーワード関連ページが閲覧者のブラウザに表示される。例えば、リンク2aがクリックされると、キーワード関連ページ3が表示される。

【0016】

この検索システムを用いた場合、WebサーバのWebアクセスログには、検索キーワードとその入力日時などを含む付加情報が記録される。

【0017】

したがって、このWebアクセスログを用いて、検索キーワードに関する分析(以下、「検索キーワード分析」という)が可能となる。

【0018】

検索キーワード分析においては、例えば、ある検索キーワードが何回入力されたか調査される。

【0019】

また、例えば、検索結果ページにおいて検索結果が実際にクリックされたか調査される。

【0020】

また、例えば、ある検索キーワードに対して、キーワード関連ページが何件抽出されたか調査される。

【0021】**【発明が解決しようとする課題】**

閲覧者は、検索システムに対して自己の意思に基づいて自由に検索キーワードを入力する。したがって、各種の検索キーワードの入力頻度は、各種の検索キーワード間で一定の割合となるのではなく、ある程度隔たりが発生する。

【0022】

このため、顧客サイトをアクセスするために入力された検索キーワードによる検索結果を量的に評価する量的評価値については、検索キーワード毎に差が生じる。

【0023】

この量的評価値の例としては、各検索キーワードについて検索結果のニーズの量、数、多さを表すニーズ値があり、例えば検索キーワードの入力回数をニーズ値とすることができます。その他にも、量的評価値の例としては、上述した検索キーワードについて抽出されたキーワード関連ページの件数、検索結果ページのクリック回数などがある。

【0024】

上述したように、一般的に検索システムは、閲覧者が入力した検索キーワードに応じて、この検索キーワードと関連が深いと判断したキーワード関連ページについての情報を抽出し、閲覧者に表示する。

【0025】

しかしながら、リスト化されているキーワード関連ページが必ずしも閲覧者の求めている内容とは限らない。

【0026】

このため、顧客サイトをアクセスするために入力された検索キーワードによる検索結果を質的に評価する質的評価値についても、検索システムの性能、リストとして表示されたキーワード関連ページの質などに応じて、検索キーワード毎に差が生じる。

【0027】

この質的評価値の例としては、検索キーワードによる検索結果がどの程度効果

的に機能したかを表す度合いである効果度がある。効果度の具体例としては、検索システムを利用した閲覧者が、その検索キーワードに対して検索システムが提示した情報サービスにどの程度満足したかを表す満足度がある。また、効果度の他の具体例として、Webサイト運営者がアクセスして欲しいと考えているWebページにどの程度自サイトの閲覧者がアクセスしているかを表す成功度がある。

【0028】

質的評価値として、例えば、閲覧者が検索キーワードによる検索結果をどの程度重視しているかを表す重視度もある。

【0029】

質的評価値を調査することは、Webページ又はWebサイトの運営者にとって、自己の運営しているWebページ又はWebサイト、検索システムが閲覧者を例えどどの程度満足させているか把握するために有効であり、Webページ又はWebサイトのデザインやコンテンツ、検索システムの改善に有効である。

【0030】

しかしながら、以下に述べる事情により、従来の技術では十分な精度で質的評価値を分析することができていなかった。

【0031】

第1に、例えば特開2002-123517号などのような従来の検索キーワード分析技術、商用ツール、非商用ツール、サービスにおいては、量的評価値のみが分析されており、例えば満足度などのような質的評価値の分析は行われていない。

【0032】

第2に、ある検索キーワードに対して、キーワード関連ページが何件抽出されたか調査することによって、この検索キーワードによるキーワード関連ページの抽出数が得られる。

【0033】

しかしながら、このキーワード関連ページの抽出数は、閲覧者の満足度などのような質的評価値を直接表すわけではない。

【0034】

例えば、閲覧者の希望するWebページの抽出漏れを防止する観点から、キーワード関連ページの抽出数が多いほど満足度が高く、少ないほど満足度が低いと定義したとする。しかしながら、実際には、キーワード関連ページの抽出数が少ない方が閲覧者にとってはリンクを選択しやすく、閲覧者の希望に合致する場合がある。

【0035】

すなわち、閲覧者の希望と合致するキーワード関連ページが検索結果として閲覧者に提示される場合、閲覧者は、冗長な検索結果が提示される場合よりもキーワード関連ページの抽出数が少ない程、検索結果に対して満足する。

【0036】

第3に、検索結果ページ上のキーワード関連ページへのリンクが実際にクリックされたか、クリックされている場合にはどのキーワード関連ページへのリンクがクリックされたか調査すると、閲覧者によって検索結果ページ上でどのキーワード関連ページが選択されたか知ることができる。

【0037】

しかしながら、アクセスされたキーワード関連ページが本当に閲覧者の求めた内容を掲載しており閲覧者が満足したか判断すること（検索結果に対して質的な評価をすること）は極めて困難である。

【0038】

本発明は、以上のような実情に鑑みてなされたもので、検索結果ページからの閲覧者の閲覧行動（2つ以上のページ遷移を含む）を分析することにより、閲覧者から入力された検索キーワードに基づく検索結果に対する質的評価値を高精度に求める検索キーワード分析プログラム及びシステム並びに方法を提供することを目的とする。

【0039】

【課題を解決するための手段】

本発明を実現するにあたって講じた具体的手段について以下に説明する。

【0040】

第1の発明の検索キーワード分析プログラムは、コンピュータに、WWWにおけるアクセス要求側を示すアクセス要求側識別データとアクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録するアクセスシーケンス抽出機能と、WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データとアクセス要求側から検索システムに入力された検索キーワードとが関連付けされている場合に、検索キーワードを抽出し、検索キーワードの抽出結果を示すキーワードデータを記録手段に記録するキーワード抽出機能と、WWWにおけるアクセスの状態から検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する質的評価値を算出し、記録手段に記録されているキーワードデータの示す検索キーワードのうちアクセスシーケンスにおいて入力された検索キーワードと質的評価値との対応関係を示す関係データを記録手段に記録する評価値算出機能とを実現させる。

【0041】

これにより、入力された検索キーワードに基づく検索システムの検索結果に対する質的評価値を容易に把握できる。この第1の発明では、アクセスシーケンス抽出機能とキーワード抽出機能とを並列に動作させることができる。

【0042】

なお、第1の発明において、評価値算出機能は、質的評価値に加えて、検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する量的評価値を算出し、記録手段に記録されているキーワードデータの示す検索キーワードのうちアクセスシーケンスにおいて入力された検索キーワードと質的評価値と量的評価値の対応関係を示す関係データを記録手段に記録するとしてもよい。

【0043】

これにより、入力された検索キーワードに基づく検索システムの検索結果に対する質的評価値、量的評価値、質的評価値と量的評価値との関係を容易に把握できる。

【0044】

第2の発明の検索キーワード分析プログラムは、コンピュータに、WWWにおけるアクセス要求側を示すアクセス要求側識別データとアクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録するアクセスシーケンス抽出機能と、WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データとアクセス要求側から検索システムに入力された検索キーワードとが関連付けられており記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスにおいて検索キーワードが入力されている場合に、検索キーワードを抽出し、検索キーワードの抽出結果を示すキーワードデータを記録手段に記録するキーワード抽出機能と、WWWにおけるアクセスの状態から検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する質的評価値を算出し、記録手段に記録されているキーワードデータの示す検索キーワードと質的評価値との対応関係を示す関係データを記録手段に記録する評価値算出機能とを実現させる。

【0045】

この第2の発明では、上記第1の発明と同様の効果を得ることができるが、アクセスシーケンス抽出機能とキーワード抽出機能とが直列に動作する特徴を持つ。

【0046】

なお、第2の発明において、評価値算出機能は、質的評価値に加えて、検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、記録手段に記録されたアクセスシーケンスデータの示すアクセスシ

ケンスに対応する量的評価値を算出し、記録手段に記録されているキーワードデータの示す検索キーワードと質的評価値と量的評価値の対応関係を示す関係データを前記記録手段に記録するとしてもよい。

【0047】

上記第1又は第2の発明の検索キーワード分析プログラムにおいて、検索キーワードと質的評価値と量的評価値の対応関係を示す関係データが求められる場合、コンピュータに、記録手段に記録された関係データにしたがって、質的評価値と量的評価値とに基づくグラフを描画するグラフ化機能を実現させるためのプログラム、を付加してもよい。これにより、分析結果の把握が一層容易になる。

【0048】

上記第1又は第2の発明の検索キーワード分析プログラムにおいて、検索キーワードと質的評価値と量的評価値とが求められる場合、量的評価値は、検索キーワードの入力数が多いほど値が大きくなる検索結果のニーズ値であり、量的評価規則は、検索キーワードの入力数が多いほどニーズ値が増加する規則であるとしてもよい。

【0049】

上記第1又は第2の発明の検索キーワード分析プログラムにおいて、Webアクセスログは、WWWにおけるアクセス毎に、アクセス要求側識別データとアクセス先識別データとアクセス時刻を示すデータとを関係付けており、質的評価規則は、アクセスシーケンスにおける検索キーワード入力後のアクセスから最後のアクセスまでの時間間隔が長いほどアクセス要求側の満足度が増加する規則であり、質的評価値は、満足度であるとしてもよい。

【0050】

上記第1又は第2の発明の検索キーワード分析プログラムにおいて、質的評価規則は、分析候補の検索キーワードと分析候補の検索キーワードの入力後にアクセスされた場合に成功度を増加するアクセス先識別データとを関連付けたデータと記録手段に記録されたキーワードデータとに基づいて、アクセスシーケンスにおいて分析候補の検索キーワードが入力されておりその入力後分析候補の検索キーワードと関連付けされているアクセス先識別データの示すアクセス先がアクセ

スされている場合に、アクセスシーケンスに対する成功度を増加する規則であり、質的評価値は、成功度であるとしてもよい。

【0051】

この場合において、質的評価規則は、分析候補の検索キーワードが入力されてから分析候補の検索キーワードと関連付けされているアクセス先識別データの示すアクセス先がアクセスされるまでのアクセス数が少ないほど加えられる成功度を大きくするとしてもよい。

【0052】

上記第1又は第2の発明の検索キーワード分析プログラムにおいて、Webアクセスログは、アクセス先識別データの示すアクセス先がハイパーリンクを使ってアクセスされた場合に、アクセス先識別データの示すアクセス先の前段でアクセスされたアクセス先を示すレファラ情報とアクセス先識別データとを関連付けたデータを含み、アクセスシーケンス抽出機能は、レファラ情報が検索システムを示す場合に、レファラ情報に関連付けされているアクセス先識別データの示すアクセス先へのアクセスを含むアクセスシーケンスを抽出し、キーワード抽出機能は、レファラ情報が検索システムを示す場合に、レファラ情報に含まれている検索キーワードを抽出し、評価値算出機能は、アクセスシーケンスに対する検索キーワード入力後の質的評価値を算出するとしてもよい。

【0053】

これにより、アクセス先で記録されたWebアクセスログのみからであっても上記と同様の効果を得ることができる。

【0054】

上記の検索キーワード分析プログラム又はこのプログラムを記録した記録媒体を用いることによって、上記各機能を有していないコンピュータ、コンピュータシステム、サーバ、クライアントに対しても、簡単にこれらの機能を付加することができる。

【0055】

本発明で実現される機能と同様の動作を実現する手段を検索キーワード分析システムに具備することにより、同様の作用効果を得ることができる。

【0056】

本発明の検索キーワード分析プログラム又はこのプログラムを記録した記録媒体をコンピュータに適用することによって、入力された検索キーワードに基づく検索システムの検索結果に対する質的評価値を容易に把握するための検索キーワード分析方法がコンピュータ上で実施される。

【0057】**【発明の実施の形態】**

以下、図面を参照しながら本発明の実施の形態について説明する。なお、Webサイトとは少なくとも1つのWebページを含むWebページのグループを指す。

【0058】

以下の各図において同一の部分については同一の符号を付してその説明を省略し、異なる部分についてのみ詳しく説明する。

【0059】**(第1の実施の形態)**

本実施の形態においては、検索システムによる検索結果に対する質的評価値を求める検索キーワード分析システムについて説明する。なお、本実施の形態においては、質的評価値のうち満足度を求める場合について説明するが、満足度ではない他の質的評価値を求める場合も同様である。

【0060】

図1は、本実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図である。

【0061】

図2は、顧客サイトの構成の一例を示すブロック図である。

【0062】

記録装置4は、調査対象Webサイトの運営者である顧客5によって運営される顧客サイト6のWebアクセスログ7を記録する。

【0063】

顧客サイト6は、サイト内検索システム6aと、一以上のWebページP1~

P_n を持つ。

【0064】

サイト内検索システム6aは、ある特定のサイト内のWebページについての検索を行う。この図1の例では、サイト内検索システム6aは、顧客サイト6のWebページ P_1 ～ P_n についての検索を行う。

【0065】

一般的な閲覧者8のうち、顧客サイト6をアクセスした閲覧者を顧客サイト閲覧者8aとする。

【0066】

表1に、Webアクセスログ7に含まれる情報の一例を示す。

【0067】

【表 1】

表 1 Web アクセスログに含まれる情報（顧客サイト閲覧者）

ホスト名	User ID	日付	時刻	リクエスト	コード	終了バイト数	転送バイト数	ユーザエージェント
pc001.○○○.co.jp	—	2002/2/14	18:56:14	/index.html	200	400	—	MOOiiia/4.0 (compatible:MOIE...)
ws20.△△△.co.jp	—	2002/2/14	18:57:30	/cgi/login.cgi	200	700	—	MOOiiia/4.7 [△△5LLL] (X11...)
pc001.○○○.co.jp	—	2002/2/14	18:58:01	/products/index.html	200	2400	—	W△m/0.3+mm□e-p21-26+mm○e-1.5.2
pc001.○○○.co.jp	—	2002/2/14	18:58:40	/products/info/p001.pdf	200	3000	—	MOOiiia/4.7 [△△5LLL] (X11...)
ws20.△△△.co.jp	—	2002/2/14	18:59:10	/cgi/order.cgi	200	700	—	MOOiiia/4.0 (compatible:MOIE...)

Web アクセスログ 7 には、顧客サイト閲覧者 8a による顧客サイト 6 のア

セス毎に、ホスト名又はIPアドレス、顧客サイト閲覧者8aが顧客サイト6にアクセスした日付と時刻を示す情報、顧客サイト閲覧者8aの発信したリクエスト(URL)、レファラ情報、アクセスした顧客サイト閲覧者8aのID(User ID)、などから構成されるレコードが含まれる。

【0068】

なお、レファラ情報とは、ハイパーリンクを使ってあるWebページがアクセスされた場合に、このアクセスされたWebページがどのWebページからリンクを辿ってアクセスされたかを示す情報である。表1においては、顧客サイト閲覧者8aのIDとレファラ情報については、具体的な値の記述を省略している。

【0069】

本実施の形態に係る検索キーワード分析システム9は、記録媒体10に記録されている検索キーワード分析プログラム11を読み出して実行する。

【0070】

検索キーワード分析プログラム11は、システム9に読み込まれ、実行されると、ログ前処理機能12、アクセスシーケンス抽出機能13、キーワード抽出機能14、評価値算出機能37を実現する。

【0071】

また、評価値算出機能37は、満足度算出機能15、満足度・キーワード関連付け機能16を含む。

【0072】

上記各機能12～16によって作成されたデータは、記録装置38に記録される。各機能12～16は、記録装置38に記録されたデータを読み込み可能である。記録装置38としては、例えば内部メモリ等が適用される。

【0073】

ログ前処理機能12は、記録装置4に記録されたWebアクセスログ7から、検索キーワード分析システム9に必要な情報を入力する。

【0074】

そして、ログ前処理機能12は、入力した情報に基づいて、各ページアクセスに対して重複のないアクセス識別子(ViewID)を付与する。

【0075】

表2に、ログ前処理機能12によって生成された前処理データの一例を示す。

【0076】

【表2】

表2 前処理データ

ViewID	ホスト名	User ID	日付	時刻	リクエスト	情報	ユーザエージェント
View1	pc001.○○○.co.jp	—	2002/2/14	18:56:14	/index.html	—	MOOiiIa/4.0 (compatible:MOIE...)
View2	ws20.△△△.co.jp	—	2002/2/14	18:57:30	/cgi/login.cgi	—	MOOiiIa/4.7 [△△5LLL] (X11...)
View3	pc001.○○○.co.jp	—	2002/2/14	18:58:01	/products/index.html	—	w△m/0.3+m□e-p21-26+m○e-1.5.2
View4	pc001.○○○.co.jp	—	2002/2/14	18:58:40	/products/info/p001.pdf	—	MOOiiIa/4.7 [△△5LLL] (X11...)
View5	ws20.△△△.co.jp	—	2002/2/14	18:59:10	/cgi/order.cgi	—	MOOiiIa/4.0 (compatible:MOIE...)

アクセスシーケンス抽出機能13は、前処理データに基づいて、顧客サイト閲覧者8aによってアクセスされた一連のWebページの列（以下、「アクセスシ

一ケンス」という)を抽出し、各アクセスシーケンスに重複のないアクセスシーケンス識別子を付与し、アクセスシーケンス抽出データを生成する。

【0077】

アクセスシーケンスとは、同一又は同一として扱われる閲覧者、すなわち同一のアクセス要求側による連続的なアクセスをいう。

【0078】

アクセスシーケンス抽出機能13は、ホスト名又はIPアドレスとユーザエージェントとの組み合わせが一致する場合に同一のアクセス要求側として扱い、ホスト名又はIPアドレスとユーザエージェントの各組み合わせに対してアクセス要求側識別子を付与し、同一のアクセス要求側から連続的にアクセスされた一以上のWebページをアクセスシーケンスとして切り出す。

【0079】

アクセスシーケンス抽出機能13は、所定のアクセスシーケンス抽出規則にしたがって、一連のページアクセスからアクセスシーケンスを切り出し、アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを作成する。

【0080】

例えば、アクセスシーケンス抽出規則として、一般的なWebアクセスログ分析で用いられているセッションを抽出する規則を適用可能である。具体的には、ある閲覧者によるWebページのアクセスから次のWebページへのアクセスの間に30分以上の時間が経過した場合(次のWebページへの遷移がなかった場合を含む)、その間をセッションの境界(もしくはセッションの終端)とし、境界と境界との間をセッションとして抽出する。そして、このセッションをアクセスシーケンスとして用いる。

【0081】

なお、アクセスシーケンスとアクセスシーケンスとの間の境界を判定するための経過時間は、上記の30分に限定されるものではなく、自由に設定可能である。すなわち、アクセスシーケンス抽出規則は、複数のアクセスが同一のアクセス要求側から所定の時間間隔内でアクセスされている場合に複数のアクセスが同一のアクセスシーケンスに含まれると判断する規則としてもよい。

【0082】

また、アクセスシーケンスを抽出する場合において、Webアクセスログ7に記録された閲覧者のID（UserID）又はCookie情報に基づいて、アクセス要求側の同一性を判断してもよい。

【0083】

表3に、アクセスシーケンスデータの一例を示す。

【0084】

【表3】

表3 アクセスシーケンスデータ

ViewID	VisitID	VisitorID	日付	時刻	リクエスト
View1	Visit1	Visitor1	2002/2/14	18:56:14	/AAA.html
View2	Visit2	Visitor2	2002/2/14	18:57:30	/BBB.cgi
View3	Visit1	Visitor1	2002/2/14	18:58:01	/CCC.html
View4	Visit2	Visitor2	2002/2/14	18:58:40	/GGG/DDD.cgi
View5	Visit3	Visitor1	2002/2/14	19:42:10	/EEE.html

表3のアクセスシーケンスデータには、各アクセスシーケンスを識別するためのアクセスシーケンス識別子（VisitID）、ホスト名とユーザエージェントの組み合わせを識別するアクセス要求側識別子（VisitorID）などから構成されるコードが含まれる。

【0085】

キーワード抽出機能14は、ログ前処理機能12によって生成された前処理データの中から、顧客サイト閲覧者8aがサイト内検索システム6aにおいて入力した検索キーワードを抽出し、検索キーワードの抽出結果を示すキーワードデータを作成する。

【0086】

表4に、Webアクセスログ7における検索システムへのリクエストの例を示す。

【0087】

【表4】

表4 検索システムへのリクエスト

例1	http://www.○○○.co.jp/search.cgi?q=xxxx
例2	http://www.○○○.co.jp/search.cgi?q=yyyy&page=10
例3	http://www.○○○.co.jp/search.cgi?q=%E3%82%8A%E3%82%93%E3%81%94

表4における例1は、検索キーワード「xxxx」を含んだ一番基本的な検索システムへのリクエストの例である。

【0088】

例2では、検索キーワード「yyyy」と検索結果の10ページ目とがリクエストされている。

【0089】

例3では、検索キーワード「りんご」を示すコードがリクエストに含まれている。

【0090】

Webアクセスログ7の中から検索キーワードを抽出するためには、検索システムの場所及び検索キーワードを示すパラメータを表すクエリーパターンに関する情報を用いる。

【0091】

表5に、検索システムの場所とクエリーパターンに関する情報の一例を示す。

【0092】

【表5】

表5 検索システムの場所とクエリーパターン

検索システムの場所	クエリーパターン
http://www.○○○.co.jp/search.cgi	q

すなわち、キーワード抽出機能14は、前処理データから検索システムの場所に関する情報に基づいて、検索システムを利用したアクセスを抽出する。

【0093】

次に、キーワード抽出機能14は、クエリーパターンに関する情報に基づいて、検索キーワードを抽出する。

【0094】

表6に、上記表4、5の例に基づいて、キーワード抽出機能14によって抽出されたパラメータの例を示す。

【0095】

【表6】

表6 抽出されたパラメータ

パラメータ	
例1	xxxx
例2	yyyy
例3	%E3%82%8A%E3%82%93%E3%81%94

抽出されたパラメータの中には、この表6の例3に示すようにエンコードされたパラメータも存在する。

【0096】

キーワード抽出機能14は、エンコードされたパラメータについては判読可能な文字列へのデコードを行い、キーワードデータを生成する。

【0097】

表7に、キーワードデータの例を示す。

【0098】

【表7】

表7 キーワード抽出データ

ViewID	検索キーワード
View6	りんご
View21	apple
View43	banana
View56	砂糖
View81	カレイ

表7のキーワードデータには、各アクセス識別子とそのアクセス識別子に対応するアクセスにおいて入力された検索キーワードとから構成されるレコードが含まれている。

【0099】

満足度算出機能15は、アクセスシーケンス抽出機能13によって生成されたアクセスシーケンスデータに基づいて、閲覧者の満足度を算出する。

【0100】

満足度を算出するために、満足度をどのように算出するかを表す質的評価規則が定義される。

【0101】

例えば、質的評価規則として、サイト内検索システム6aに検索キーワードを入力した後の顧客サイト閲覧者8aによる顧客サイト6の閲覧時間（以下、「検索後閲覧時間」という）が長いほど満足度が高くなる規則を定義する。

【0102】

表8に、検索後閲覧時間と満足度とを関係付けた例を示す。

【0103】

【表8】

表8 検索後閲覧時間と満足度

検索後閲覧時間	満足度
閲覧なし	0
1ページのみの閲覧	0
1分以下	10
1-2分	20
2-3分	30
3-5分	40
5-7分	50
7-10分	60
10-15分	70
15-20分	80
20-25分	90
30分以上	100

検索後閲覧時間は、満足度を求めるアクセスシーケンスについて最後のアクセス時刻から最初のキーワード関連ページへのアクセス時刻を引いて得られる時間間隔とする。ただし、検索結果ページからキーワード関連ページに遷移することなくアクセスシーケンスが終わっていた場合には検索後閲覧時間を「閲覧なし」とする。また、検索結果ページからキーワード関連ページに遷移して、その後他

のサイト内のWebページを閲覧することなくアクセスシーケンスが終わっている場合には、検索後閲覧時間を「1ページのみの閲覧」とする。

【0104】

すなわち、満足度算出機能15は、アクセスシーケンスデータを読み込み、アクセスシーケンスごとに検索後閲覧時間を計算する。

【0105】

次に、満足度算出機能15は、上記表8のような検索後閲覧時間と満足度の対応データを読み込み、アクセスシーケンスごとの満足度を算出する。

【0106】

表9に、満足度算出機能15によって生成された満足度算出データの例を示す。

【0107】

【表9】

表9 満足度算出データ

ViewID	VisitID	満足度
View1	Visit1	0
View2	Visit2	20
View3	Visit1	0
View4	Visit2	20
View5	Visit3	10

表9の満足度算出データには、各アクセス識別子とそのアクセスの属するアクセスシーケンス識別子に対応する満足度とから構成されるレコードが含まれている。

【0108】

満足度・キーワード関連付け機能16は、満足度算出機能15で生成された満足度算出データと、キーワード抽出機能14で生成されたキーワードデータを統合し、検索キーワードごとの満足度を算出し、満足度と検索キーワードとを関係付け、関係データを作成する。

【0109】

満足度・キーワード関連付け機能16は、まず、満足度算出データとキーワー

ドデータを読み込む。

【0110】

次に、満足度・キーワード関連付け機能16は、アクセス識別子に基づいて、満足度算出データとキーワードデータを統合する。

【0111】

表10に、満足度算出データとキーワードデータの統合結果を示す。

【0112】

【表10】

表10 満足度算出データとキーワードデータの統合結果

ViewID	満足度	検索キーワード
View6	10	りんご
View21	0	apple
View83	30	りんご
View121	20	砂糖
View137	0	apple

そして、満足度・キーワード関連付け機能16は、検索キーワードごとに満足度の平均を求め、関係データを作成する。

【0113】

表11に、関係データの例を示す。

【0114】

【表11】

表11 関係データ

検索キーワード	満足度
りんご	23.4
apple	17.8
砂糖	2.4
肉野菜	2.4
蜂蜜	11.6

図3は、本実施の形態に係る検索キーワード分析システム9の動作の一例を示すフローチャートである。

【0115】

ステップS1において、ログ前処理機能12は、Webアクセスログ7に含まれる情報から必要な情報を入力し、前処理データを作成する。

【0116】

ステップS2において、アクセスシーケンス抽出機能13は、前処理データとアクセスシーケンス抽出規則とに基づいて、アクセスシーケンスを抽出し、アクセスシーケンスデータを作成する。

【0117】

ステップS3において、満足度算出機能15は、アクセスシーケンスデータと満足度算出規則とに基づいて、アクセスシーケンス毎の満足度を算出し、満足度算出データを作成する。

【0118】

ステップS4において、キーワード抽出機能14は、前処理データに基づいて、サイト内検索システム6aに入力された検索キーワードを抽出し、キーワードデータを作成する。

【0119】

ステップS5において、満足度・キーワード関連付け機能16は、満足度算出データと、キーワードデータに基づいて、検索キーワードごとの満足度を算出し、関係データを作成する。

【0120】

なお、ステップS4は、ステップS2及びステップS3の前に実行されてもよく、後に実行されてもよく、並列に実行されてもよい。

【0121】

図4は、本実施の形態に係るキーワード抽出機能14の動作の一例を示すフローチャートである。

【0122】

ステップT1において、キーワード抽出機能14は、検索システムの場所とクエリーパターンに関する情報を読み込む。

【0123】

ステップT2において、キーワード抽出機能14は、ログ前処理機能12によ

って前処理されたWebアクセスログ7から一つのレコードを読み込む。

【0124】

ステップT3において、キーワード抽出機能14は、読み込んだレコードに含まれているリクエストが検索システムの場所を含むか判定する。

【0125】

リクエストが検索システムの場所を含むと判定された場合、ステップT4において、キーワード抽出機能14は、検索システムの場所に対応するクエリーパターンに基づいて、リクエストからパラメータを抽出する。

【0126】

リクエストが検索システムの場所を含むと判定されなかった場合、キーワード抽出機能14は、ステップT2以降の処理を繰り返す。

【0127】

ステップT5において、キーワード抽出機能14は、抽出したパラメータに対してデコードが必要か否か判定する。

【0128】

デコードが必要な場合、ステップT6において、キーワード抽出機能14は、抽出したパラメータをデコードする。

【0129】

ステップT7において、キーワード抽出機能14は、ログ前処理機能12によって前処理されたWebアクセスログ7から終端のレコードを読み込んだか否か判定する。

【0130】

終端のレコードが読み込まれていないと判定された場合、キーワード抽出機能14は、ステップT2以降の処理を繰り返す。

【0131】

終端のレコードが読み込まれたと判定された場合、ステップT8において、キーワード抽出機能14は、抽出されたパラメータを検索キーワードとし、キーワードデータを作成する。

【0132】

図5は、本実施の形態に係る満足度算出機能15の動作の一例を示すフローチャートである。

【0133】

ステップU1において、満足度算出機能15は、アクセスシーケンス抽出機能13によって生成されたアクセスシーケンスデータを読み込む。

【0134】

ステップU2において、満足度算出機能15は、アクセスシーケンス毎に、検索後閲覧時間を計算する。

【0135】

ステップU3において、満足度算出機能15は、検索後閲覧時間と満足度とを関係付けたデータを読み込む。

【0136】

ステップU4において、満足度算出機能15は、アクセスシーケンス毎に、検索後閲覧時間に対応する満足度を求める。

【0137】

ステップU5において、満足度算出機能15は、アクセスシーケンス毎の満足度を示す満足度算出データを作成する。

【0138】

なお、ステップU3は、ステップU1及びステップU2の前に実行されてもよく、並列に実行されてもよい。

【0139】

以上説明した本実施の形態においては、検索キーワードについての分析結果として、顧客サイト閲覧者8aがその検索キーワードによる検索結果についてどれほど満足したか求めることができ、検索キーワードに対する検索結果についての質的な評価を分析することができる。

【0140】

顧客5は、検索結果に対して顧客サイト閲覧者8aがどの程度満足しているかを定量的に知ることができる。例えば、上記表11の結果より、検索キーワード「りんご」による検索結果については顧客サイト閲覧者8aの満足度が高いのに

対し、検索キーワード「砂糖」又は検索キーワード「肉野菜」による検索結果については顧客サイト閲覧者8aの満足度が低いことが分かる。

【0141】

なお、質的評価値としては、満足度等に代表される効果度の他に、例えば重視度や利便度も考えられる。

【0142】

重視度とは、Webサイトの閲覧者が検索キーワードに関連する事柄について、どの程度重視しているかを表す指標である。

【0143】

利便度とは、Webサイトの閲覧者が検索キーワードに関連する事柄について、どの程度容易に閲覧やWebサイトへの書き込み等の操作ができたかを表す指標である。

【0144】

重視度や利便度も、満足度と同様に検索後閲覧時間によって求めることが可能である。

【0145】

検索後閲覧時間を満足度、重視度、利便度のいずれを算出するのに利用するかは、分析対象となるWebサイトや閲覧者の性質、分析者の観点によって適宜選択できる。

【0146】

例えば、商品を販売しているECサイトにおいて、他にも類似の競合サイトが多く存在するケースでは、顧客は検索キーワードに関連したコンテンツが顧客の期待にそわなかつた場合に、比較的短時間で閲覧をあきらめてサイトを去り、別のサイトを閲覧することが考えられる。このようなケースでは、検索後閲覧時間は満足度としての利用が適切である。例えば、「蜂蜜」という検索キーワードに対する検索後閲覧時間が短いために、「蜂蜜に関する満足度が低い」という分析結果が得られた場合、この結果は「蜂蜜」に関連するコンテンツを改良して、満足度を向上させるような対策を打つ必要性が高い、という判断をする根拠として有効である。

【0147】

一方、他に類似の競合サイトが少ない場合や、提供しているコンテンツが閲覧者の嗜好を満たすような内容というよりはむしろ合理的必要性に応じて閲覧している場合や、その検索キーワードに関する情報をできるだけ当該Webサイトから得たいと思っている閲覧者が多数であるような場合においては、閲覧者が検索キーワードに関連する内容を重視している度合いに応じて、長時間サイトに滞在して、その情報を閲覧すると考えられる。したがって、検索後閲覧時間を重視度の算出に利用することが適切な場合がある。この場合に、ある検索キーワードに対する検索後閲覧時間が短いために、それに関する重視度が低いという分析結果が得られた場合、この結果はそのコンテンツについては以後それほど改良のための対策を打つ必要性が低い、という判断をする根拠として有効である。

【0148】

さらに、利便度においては、例えば提供しているコンテンツをどれだけ閲覧者が容易に理解したか、情報の入力がどの程度容易であったかを評価する指標として有効である。

【0149】

例えば検索キーワードに応じてアンケート内容を変えているWebサイトにおいて、検索後閲覧時間が長いということは、アンケートの返答の入力に手間取っていると判断され、利便度が低いと判断される。

【0150】

このように利便度の低いコンテンツがあれば、利便度を向上させるべく、対策を打つ必要性が高い、という判断をする根拠として有効である。

【0151】

（第2の実施の形態）

本実施の形態においては、上記第1の実施の形態に係る検索キーワード分析システムの変形例について説明する。なお、本実施の形態においては、質的評価値のうち満足度を求め、量的評価値のうちニーズ値を求める場合について説明するが、他の質的評価値及び量的評価値を求める場合も同様である。

【0152】

図6は、本実施の形態に係る検索キーワード分析システムの構成の一例を示す
ブロック図である。

【0153】

本実施の形態に係る検索キーワード分析システム18は、記録媒体19に記録
されている検索キーワード分析プログラム20を読み出して実行する。

【0154】

検索キーワード分析プログラム20は、システム18に読み込まれ、実行され
ると、ログ前処理機能12、アクセスシーケンス抽出機能13、キーワード抽出
機能14、評価値算出機能39を実現する。

【0155】

また、評価値算出機能39は、満足度算出機能15、ニーズ値算出機能21、
満足度・ニーズ値関連付け機能22を含む。

【0156】

ニーズ値算出機能21、満足度・ニーズ値関連付け機能22によって作成され
たデータは、記録装置38に記録される。ニーズ値算出機能21、満足度・ニー
ズ値関連付け機能22は、記録装置38に記録されたデータを読み込み可能であ
る。

【0157】

ニーズ値算出機能21は、キーワード抽出機能14によって生成されたキーワ
ードデータに基づいて、検索キーワードごとのニーズ値を算出する。

【0158】

ニーズ値を算出するために、ニーズ値をどのように算出するかを表す量的評価
規則が定義される。

【0159】

例えば、量的評価規則として、検索キーワードの入力総数に対する各検索キ
ーワードの占める割合を百分率で表した値をニーズ値と定義する。

【0160】

ニーズ値算出機能21は、キーワード抽出機能14によって生成されたキーワ
ードデータを読み込み、検索キーワードごとの入力個数をカウントする。

【0161】

次に、ニーズ値算出機能21は、検索キーワードごとの入力個数から全検索キーワードの入力個数の合計を算出する。

【0162】

次に、ニーズ値算出機能21は、全検索キーワードの入力個数と検索キーワードごとの入力個数に基づいて、全検索キーワードの入力に対する各検索キーワードの占める割合を百分率で表す。

【0163】

そして、ニーズ値算出機能21は、検索キーワードごとのニーズ値を表すニーズ値算出データを作成する。

【0164】

表12に、ニーズ値算出データの一例を示す。

【0165】

【表12】

表12 ニーズ値算出データ

検索キーワード	ニーズ値
りんご	21.5
apple	14.6
砂糖	3.6
肉野菜	17.6
蜂蜜	2.7

満足度・ニーズ値関連付け機能22は、満足度算出機能15によって生成された満足度算出データとニーズ値算出機能21によって生成されたニーズ値算出データを読み込む。

【0166】

そして、満足度・ニーズ値関連付け機能22は、満足度算出データとニーズ値算出データを検索キーワード毎に統合し、検索キーワードごとのニーズ値、満足度を表す関係データを作成する。

【0167】

表13に、関係データの一例を示す。

【0168】

【表13】

表13 関係データ

検索キーワード	ニーズ値	満足度
りんご	21.5	23.4
apple	14.6	17.8
砂糖	3.6	2.4
肉野菜	17.6	2.4
蜂蜜	2.7	11.6

以上説明した本実施の形態において、顧客5は、質的評価値（例えば満足度）と量的評価値（例えばニーズ値）とを参照、比較しつつ、早急に対策を取る必要がある検索キーワードについて知ることができる。

【0169】

例えば、顧客5は、上記表13の例を参照した場合に、検索キーワード「砂糖」と「肉野菜」は同程度の満足度であり他の検索キーワードに比べて満足度が低いが、「肉野菜」は「砂糖」に比べてニーズ値が高いと解釈できる。

【0170】

この結果、顧客5は、検索キーワード「砂糖」について満足度が低いがあまり検索のニーズがなく（本実施形態の場合、検索キーワード「砂糖」による検索結果ページの閲覧者が少なく）、比較的影響力が低いと判断できる。

【0171】

これに対し、顧客5は、検索キーワード「肉野菜」は「砂糖」と同じ程度で満足度が低いが検索のニーズがあり（本実施形態の場合、検索キーワード「肉野菜」による検索結果ページの閲覧者が多く）、影響力が大きいと判断できる。

【0172】

このため、本実施の形態において、顧客5は、「肉野菜」関連のコンテンツを増やすなどの対策を優先的に講じる必要があると判断でき、迅速かつ正確に閲覧者の希望に合うようにサイト内検索システム6a、WebページP₁～P_nを変更することができる。

【0173】

なお、本実施の形態において、ニーズ値の算出方法は、キーワード抽出機能14、ニーズ値算出機能21の設定を変更することにより、自由に変更できる。例えば、検索キーワードが入力されて得られたキーワード関連ページをアクセスした顧客サイト閲覧者8aの人数又はキーワード関連ページのアクセス数に基づいてニーズ値を算出してもよい。また、例えば、アクセスシーケンスにしたがったアクセスを行った閲覧者の人数、アクセスシーケンスにしたがったアクセスの実行された回数（閲覧回数）に基づいてニーズ値を算出してもよい。

【0174】

なお、量的評価値の例として、ニーズ値の他に機会値もある。機会値とは、Webサイトがその目的を遂行するに当たって、検索キーワードに関連する観点で、どれだけの事業機会を得たかを示す指標である。

【0175】

例えば、食品を販売しているECサイトにおいて、「砂糖」という検索キーワードによって閲覧してきた閲覧者が多数いれば、それだけ多くの砂糖を販売する事業機会を得たことになる。

【0176】

また、検索キーワードごとの入力回数のような、ニーズ値の算出に用いられた数量に基づいて機会値を算出してもよい。

【0177】

検索キーワードごとの入力回数をニーズ値とするか機会値とするかについては、Webサイトや閲覧者の特性、分析者の観点に依存する。

【0178】

例えば、商品そのものを販売するというよりは、商品に関する情報を提供し、その反応を知るために運営されているWebサイトにおいては、検索キーワードごとの入力回数をニーズ値として取り扱うことが適切である。

【0179】

一方、商品そのものをオンラインで販売することを主たる目的としているWebサイトでは検索キーワードごとの入力回数を機会値として取り扱い、各検索キーワードに関連した商品の販売の機会の量を示す指標とすることが適切である。

【0180】

(第3の実施の形態)

本実施の形態においては、上記第2の実施の形態に係る検索キーワード分析システムの変形例について説明する。

【0181】

図7は、本実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図である。

【0182】

本実施の形態に係る検索キーワード分析システム24は、記録媒体25に記録されている検索キーワード分析プログラム26を読み出して実行する。

【0183】

検索キーワード分析プログラム26は、システム24に読み込まれ、実行されると、ログ前処理機能12、アクセスシーケンス抽出機能13、キーワード抽出機能14、評価値算出機能39に加えて、グラフ化機能27を実現する。

【0184】

グラフ化機能27は、満足度・ニーズ値関連付け機能22によって生成された関係データを読み込む。

【0185】

そして、グラフ化機能27は、検索キーワードごとのニーズ値・満足度の値をグラフで表現した出力データ28を作成する。

【0186】

グラフ化機能27は、所定のグラフ化規則とグラフ化に利用されるパラメータとにしたがって、検索キーワードごとに満足度とニーズ値とが関連付けされた関係データをグラフ化する。

【0187】

例えば、グラフ化機能27は、散布図を作成する。この場合、満足度とニーズ値とをパラメータとし、散布図の一方の軸（X軸）をニーズ値、他方の軸（Y軸）を満足度とし、関係データにおけるニーズ値・満足度の値に基づいて各検索キーワードを表す点をプロットする。

【0188】

図8は、検索キーワード毎の満足度とニーズ値とを表すグラフの一例である。

【0189】

本実施の形態においては、検索キーワードごとの満足度とニーズ値とがグラフ化されて表示されるため、顧客5に視覚的な分析結果が提供される。

【0190】

したがって、顧客5は、より人間が直感的に解釈・判断しやすい分析結果を得ることができる。

【0191】

例えば、顧客5は、上記図8のグラフから、検索キーワード「りんご」についてはニーズ値が高く満足度も高いのに対して、検索キーワード「スーパー」「肉野菜」についてはニーズ値が高いにもかかわらず満足度が低く、早急にサイト内検索システム6a、WebページP₁～P_nを変更し、対策を取る必要があると解釈される。

【0192】

なお、本実施の形態におけるグラフ化機能27の設定を変更することにより、グラフの軸の値を変換でき、グラフ表示効果を向上させることができる。例えば、グラフ化機能27は、ニーズ値、満足度を0から1の間に正規化するとしてもよい。また、ニーズ値、満足度に対数処理を施してグラフ表示してもよい。さらに、グラフ化機能27は、ニーズ値、満足度を棒グラフで表してもよい。さらに、グラフ化機能27は、グラフに適切な補助線を加えてもよい。

【0193】**(第4の実施の形態)**

本実施の形態においては、上記第1乃至第3の実施の形態に係る検索キーワード分析システムの変形例について説明する。なお、以下においては、上記第1の実施の形態についての変形例を説明するが、他の第2、第3の実施の形態についても同様の変形が可能である。

【0194】

なお、本実施の形態においては、質的評価値のうち成功度を求める場合につい

て説明するが、成功度ではない他の質的評価値を求める場合も同様である。

【0195】

図9は、本実施の形態に係る検索キーワード分析システムの一例を示すブロック図である。

【0196】

本実施の形態に係る検索キーワード分析システム29は、記録媒体30に記録されている検索キーワード分析プログラム31を読み出して実行する。

【0197】

検索キーワード分析プログラム31は、システム29に読み込まれ、実行されると、ログ前処理機能12、アクセスシーケンス抽出機能13、キーワード抽出機能14、評価値算出機能40を実現する。

【0198】

また、評価値算出機能40は、成功度算出機能32、成功度・キーワード関連付け機能41を含む。

【0199】

成功度算出機能32は、質的評価規則にしたがって、アクセスシーケンス抽出機能13によって算出されたアクセスシーケンスデータとキーワード抽出機能14によって算出されたキーワードデータに基づいて、成功度を算出する。

【0200】

本実施の形態に係る質的評価規則として、例えば以下のような内容の規則が定義される。

【0201】

顧客5は、代表的な検索キーワード（分析候補の検索キーワード）に対して、閲覧者にアクセスして欲しいWebページ（以下、「アクセス希望ページ」という）の識別情報を関連付けたリスト（以下、「アクセス希望ページリスト」という）を作成する。

【0202】

そして、顧客サイト閲覧者8aが検索結果ページ上のリンクにしたがってアクセス希望ページをアクセスしたら成功度をカウントアップする。

【0203】

ただし、検索結果ページからアクセス希望ページまでのリンククリック数が増加するにつれて成功度を減少させる。

【0204】

なお、顧客5が指定しなかった検索キーワードは「（その他）」とし、成功度は一切加算せず評価対象外とする。

【0205】

以上述べた質的評価規則において利用されるアクセス希望ページリストの一例を、表14に示す。

【0206】

【表14】

表14 アクセス希望ページリスト

検索キーワード	アクセス希望ページ
りんご	/apple/index.html
apple	/apple/index.html
肉野菜	/meat/index.html
肉野菜	/vegetable/index.html
...	...
(その他)	

なお、一つの検索キーワードに対して複数のアクセス希望ページが対応付けされていてもよく、複数の検索キーワードが一つのアクセス希望ページに対応付けされていてもよい。

【0207】

また、表15に、検索結果ページからアクセス希望ページまでのリンククリック数が増加するにつれて成功度を減少させるための成功度重み付け表の一例を示す。

【0208】

【表15】

表15 成功度重み付け表

検索結果ページ からのクリック数	成功度
1	100
2	80
3	60
4	40
5 以上	20
到達せず	0

成功度算出機能32は、アクセスシーケンス抽出機能13によって生成されたアクセスシーケンスデータと、キーワード抽出機能14によって生成されたキーワードデータと、アクセス希望ページリストと、成功度重み付け表とを読み込む。

【0209】

なお、成功度算出機能32は、キーワードデータの全てを読み込むのではなく、キーワードデータのうち検索キーワードとアクセス識別子の組み合わせのみを読み込むとしてもよい。

【0210】

次に、成功度算出機能32は、キーワードデータから、アクセス希望ページリストの検索キーワードに含まれている検索キーワードとそのアクセス識別子(ViewID)を取り出す。

【0211】

次に、成功度算出機能32は、アクセスシーケンスデータにおいて、上記作業によって取り出されたアクセス識別子と関連付けされているアクセスシーケンス識別子を取り出す。

【0212】

次に、成功度算出機能32は、取り出されたアクセスシーケンス識別子に関連するアクセス識別子とリクエストとをアクセスシーケンスデータから取り出し、リクエストに含まれている検索キーワードを取り出す。

【0213】

表16に、この段階における第1の中間生成データを示す。

【0214】

【表16】

表16 中間生成データ (1)

ViewID	VisitID	リクエスト	検索キーワード
View20	Visit6	/index.html	
View21	Visit6	/search.cgi	
View22	Visit6	/search.cgi?q=%E3%82%A%E3%82%93%E3%81%94	りんご
View26	Visit6	/general/index.html	
View28	Visit6	/apple/index.html	
View37	Visit6	/apple/tsugaru.html	
View45	Visit7	/index.html	
View47	Visit7	/search.cgi	
View48	Visit7	/search.cgi?q=apple	apple
...

次に、成功度算出機能32は、各検索キーワードに対するアクセス希望ページがその検索キーワードが入力されたアクセスシーケンスにおいてアクセスされているかを調査する。

【0215】

アクセス希望ページがアクセスされている場合、成功度算出機能32は、検索キーワードの入力から何クリック後にアクセス希望ページがアクセスされたかを調査する。

【0216】

なお、検索キーワードの入力前にアクセス希望ページがアクセスされていても、そのアクセスは成功度の算出において考慮しない。

【0217】

上記表16の例において、アクセス識別子「View22」に関連する検索キーワード「りんご」に対するアクセス希望ページ識別子は、「/apple/index.html」である。このアクセス希望ページ識別子「/apple/index.html」は、アクセス識別子「View22」と同一のアクセスシーケンス識別子「Visit6」と関連するアクセス識別子「View28」に関連付けされている。

【0218】

そして、アクセス希望ページ識別子「/apple/index.html」は、検索キーワード「りんご」が入力された2クリック後に、入力されている。

【0219】

次に、成功度算出機能32は、成功度と検索結果ページからのクリック数との関係に基づいて、成功度を算出する。

【0220】

上記の例及び上記表15を適用した場合、アクセスシーケンス識別子「Visit6」で表されるアクセスシーケンスでは、検索キーワード「りんご」に対して満足度「80」が与えられる。

【0221】

表17に、この段階における第2の中間生成データを示す。

【0222】

【表17】

表17 中間生成データ（2）

VisitID	検索キーワード	成功度
Visit6	りんご	80
Visit7	apple	100
Visit21	肉野菜	0
Visit23	スーパー	60
Visit30	りんご	20

そして、成功度算出機能32は、検索キーワードごとに、成功度の平均値を求めた成功度算出データを作成する。

【0223】

成功度・キーワード関連付け機能41は、成功度算出機能32で生成された成功度算出データと、キーワード抽出機能14で生成されたキーワードデータを統合し、検索キーワードごとの成功度を算出し、成功度と検索キーワードとを関係付け、関係データを作成する。

【0224】

関係データは、例えば、項目「満足度」が項目「成功度」に変化する点を除い

て上記表11と同様の形式となる。

【0225】

以上説明した本実施の形態における成功度算出方法を用いることで、ある検索キーワードに対し顧客にアクセスしてもらいたいとと考えているWebページへ適切に誘導がなされているかを知ることができる。

【0226】

(第5の実施の形態)

本実施の形態においては、上記第1から第4の実施の形態に係る検索キーワード分析システムの変形例について説明する。

【0227】

本実施の形態においては、検索システムが顧客サイト外にある場合の検索キーワード分析システムについて説明する。

【0228】

なお、以下においては、上記第1の実施の形態についての変形例を説明するが、他の第2から第4の実施の形態についても同様の変形が可能である。

【0229】

図10は、本実施の形態に係る顧客サイトの構成の一例を示す図である。

【0230】

顧客サイト61は一以上のWebページP₁～P_nを持つ。また、顧客サイト61におけるいくつかのWebページは、顧客5以外の第3の運営主体によって運営されている複数のサイト外の検索システム（以下、「サイト外検索システム」という）341～34nの検索対象となっている。

【0231】

サイト外検索システム341～34nは、検索キーワードを入力すると、キーワード関連ページへのリンクを含む検索結果ページを表示する。

【0232】

本実施の形態においては、顧客サイト61と検索システム341～34nの運営主体が異なるため、上記第1から第4の実施形態の場合と異なり検索システム341～34nのWebアクセスログを取得することは困難である。

【0233】

したがって、本実施の形態においては、検索システム341～34nのWebアクセスログではなく、顧客サイト61のWebアクセスログに含まれるレファラ情報をを利用して検索キーワード分析を行う。

【0234】

表18に、顧客サイト61のWebアクセスログの一例を示す。

【0235】

【表18】

表18 レファラ情報付Webアクセスログ

ホスト名	User ID	日付	時刻	リクエスト	終了コード	転送バイト数	レファラ情報	ユーザエージェント
pc001.○○○.	—	2002/2/14	18:56:14	/index.html	200	400	○ ne.jp/□□□...	MOOiiia/4.0 (compatible:MOIE...)
ws20.△△△.	co.jp	—	2002/2/14	18:57:30	/cgi/login.cgi	200	700	—
pc001.○○○.	co.jp	—	2002/2/14	18:58:01	/products/index.html	200	2400	http://www.×× X. com/△△△...
pc001.○○○.	co.jp	—	2002/2/14	18:58:40	/products/info/p001.pdf	200	3000	http://www.○○ ○ ne.jp/○○○...
ws20.△△△.	co.jp	—	2002/2/14	18:59:10	/cgi/order.cgi	200	700	http://www.○○ ○ ne.jp/□□□...
								(compatible:MOIE...)

図11は、サイト外検索システム341から顧客サイト61への閲覧遷移の一

例を示すブロック図である。

【0236】

閲覧者によって検索キーワード入力画面35に検索キーワードが入力されると、検索結果ページ36が作成される。この検索結果ページ36を示すURLには、検索キーワード入力画面35に入力された検索キーワードが含まれる場合がある。

【0237】

閲覧者によって検索結果ページ36のリンクがクリックされ、顧客サイト61のキーワード関連ページが閲覧された場合、検索キーワードを含む検索結果ページのURLであるレファラ情報とアクセスされた顧客サイト61のキーワード関連ページのURLであるリクエストとが関連付けされて顧客サイト61のWebアクセスログに記録される。

【0238】

本実施の形態に係る検索キーワード分析システムは、上記図1に示す構成と同様の構成を持つため、ここでは図1を参照して本実施の形態の特徴的な部分について説明する。

【0239】

本実施の形態に係る顧客サイト61のWebアクセスログには、レファラ情報が含まれる。

【0240】

本実施の形態に係る検索キーワード分析システムにおいては、アクセスシーケンス抽出機能、キーワード抽出機能、満足度抽出機能の動作が、上記第1の実施の形態に係る検索キーワード分析システム9と異なる。

【0241】

本実施の形態に係るアクセスシーケンス抽出機能は、アクセスシーケンス識別子の付与、アクセスシーケンスの区切りの判定などについては、上記第1の実施の形態に係るアクセスシーケンス抽出機能13と同様の処理を実行する。

【0242】

また、本実施の形態に係るアクセスシーケンス抽出機能は、各アクセスについ

てのレファラ情報に基づいて、レファラ情報がサイト外検索システムを示すか否か判定する。

【0243】

表19に、サイト外検索システムを表すURLパターンと検索キーワードを抽出するためのクエリーパターンとを関係付けたデータを示す。

【0244】

【表19】

表19 パターン

検索システムのURL	クエリーパターン
http://www.△△△.com/	keyword
http://www.□□□.ne.jp/	kw
http://www.○○○.co.jp/	query
http://www.×××.co.jp/	word
http://www.○○○.com/	q

本実施の形態に係るアクセスシーケンス抽出機能は、サイト外検索システムからのアクセスか否かを判定する場合に、各アクセスのレファラ情報が検索システムのURLパターンを含むか調査する。

【0245】

そして、本実施の形態に係るアクセスシーケンス抽出機能は、レファラ情報がサイト外検索システムを示すアクセスに対してレファラフラグを付与したアクセスシーケンスデータを作成する。

【0246】

表20に、本実施の形態に係るアクセスシーケンスデータの一例を示す。

【0247】

【表20】

表20 アクセスシーケンスデータ

ViewID	VisitID	VisitorID	日付	時刻	リクエスト	レファラ フラグ
View1	Visit1	Visitor1	2002/2/14	18:56:14	/AAA.html	1
View2	Visit2	Visitor2	2002/2/14	18:57:30	/BBB.cgi	0
View3	Visit1	Visitor1	2002/2/14	18:58:01	/CCC.html	0
View4	Visit2	Visitor2	2002/2/14	18:58:40	/GGG/DDD.cgi	0
View5	Visit3	Visitor1	2002/2/14	19:42:10	/EEE.html	1

表20において、レファラフラグ「0」は、サイト外検索システムからのアクセスでない旨を示す。

【0248】

一方、レファラフラグ「1」は、サイト外検索システムからのアクセスである旨を示す。

【0249】

このレファラフラグを記録することにより、アクセスシーケンスの途中で検索が行われた場合であっても、検索後のアクセス状況を調査し、満足度を算出できる。

【0250】

本実施の形態に係るキーワード抽出機能は、ログ前処理機能12によって生成された前処理データに基づいて、レファラ情報が検索システムのURLパターンを含むアクセスを求める。

【0251】

そして、本実施の形態に係るキーワード抽出機能は、URLパターンが含まれているレファラ情報の中から、その検索システムのURLパターンに対応付けられたクエリーパターンに基づいて、閲覧者がサイト外検索システムにおいて入力した検索キーワードを抽出する。

【0252】

表21に、本実施の形態に係るキーワード抽出機能によって生成されたキーワードデータの一例を示す。

【0253】

【表21】

表21 キーワードデータ

ViewID	レファラ情報に含まれる キーワード
View6	りんご
View21	apple
View43	banana
View56	砂糖
View81	カレイ

本実施の形態に係る満足度算出機能は、アクセスシーケンス抽出機能によって生成されたアクセスシーケンスデータから満足度を算出する。

【0254】

満足度の算出方法としては、例えば上記第1の実施の形態で説明した規則を適用する。

【0255】

すなわち、本実施の形態に係る満足度算出機能は、本実施の形態に係るアクセスシーケンス抽出機能によって生成されたアクセスシーケンスデータを読み込んだ後、Webページのアクセスごとに検索後閲覧時間を計算する。

【0256】

検索後閲覧時間は、同一のアクセスシーケンス識別子で表されるアクセスの中で、検索システムを利用したレファラ情報を持つアクセス（上記表20においてフラグ「1」のアクセス）から最後に閲覧されたアクセスまでの時刻の差分を求ることで計算できる。

【0257】

なお、検索システムをレファラ情報として持つアクセスがないアクセスシーケンスに関しては検索後閲覧時間を記録しない。

【0258】

また、本実施の形態に係る満足度算出機能は、上記表8のような検索後閲覧時間と満足度とを関連付けたデータを読み込み、アクセスごとの満足度を算出する

。

【0259】

ただし、検索後閲覧時間の記録がない場合は満足度を記録しない。

【0260】

そして、本実施の形態に係る満足度算出機能は、アクセス識別子毎にアクセスシーケンス識別子と満足度とを関連付け、上記表9と同様な満足度算出データを作成する。

【0261】

以上説明した本実施の形態においては、顧客サイト61が検索システムを具備していない場合であっても、検索キーワードの満足度を算出することができる。

【0262】

したがって、検索キーワードの分析結果を得ることができる顧客サイトを大幅に増加することができる。

【0263】

(第6の実施の形態)

本実施の形態においては、第1の実施の形態に係る検索キーワード分析システムの変形例について説明する。

【0264】

第1から第5の実施の形態においては、アクセスシーケンスの抽出や質的評価値の算出と、検索キーワードの抽出を並列に実行し、最終的に関連付け機能によって統合する方法を用いていた。

【0265】

これに対し、本実施の形態では、検索キーワードの抽出処理と質的評価値の算出処理を直列に行う。

【0266】

なお、本実施の形態では、サイト内検索について分析を行う場合について記述しているが、本実施の形態と第5の実施形態とを組み合わせることでサイト外検索についての分析にも適用可能である。

【0267】

図12は、本実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図である。

【0268】

本実施の形態に係る検索キーワード分析システム42は、記録媒体43に記録されている検索キーワード分析プログラム44を読み出して実行する。

【0269】

検索キーワード分析プログラム44は、システム42に読み込まれ、実行されると、ログ前処理機能12、アクセスシーケンス抽出機能13、キーワード抽出機能45、評価値算出機能46を実現する。

【0270】

キーワード抽出機能45、評価値算出機能46によって作成されたデータは、記録装置38に記録される。キーワード抽出機能45、評価値算出機能46は、記録装置38に記録されたデータを読み込み可能である。

【0271】

キーワード抽出機能45は、アクセスシーケンス抽出機能13によって生成されたデータから、顧客サイト閲覧者8aがサイト内検索システム6aにおいて入力した検索キーワードを抽出する。検索キーワードの抽出には上記第1の実施の形態と同等の方法を用いる。

【0272】

すなわち、キーワード抽出機能45は、前処理データから検索システムの場所に関する情報に基づいて、検索システムに対するアクセスを抽出する。

【0273】

次に、キーワード抽出機能45は、クエリーパターンに関する情報に基づいて、検索キーワードを抽出する。なお、エンコードされたパラメータについては判読可能な文字列へのデコードを行う。

【0274】

そして、キーワード抽出機能45は、アクセスシーケンス・キーワードデータを生成する。この際、検索システムを利用していないアクセスに関してはキーワード項目に何も記録しない。

【0275】

表22に、アクセスシーケンス・キーワードデータの具体例を示す。

【0276】

【表22】

表22 アクセスシーケンス・キーワードデータ

ViewID	VisitID	VisitorID	日付	時刻	検索キーワード
View1	Visit1	Visitor1	2002/2/14	18:56:14	
View2	Visit2	Visitor2	2002/2/14	18:57:30	りんご
View3	Visit1	Visitor1	2002/2/14	18:58:01	
View4	Visit2	Visitor2	2002/2/14	18:58:40	
View5	Visit3	Visitor1	2002/2/14	19:42:10	たまねぎ

評価値算出機能46は、アクセスシーケンス・キーワードデータに基づいて、質的評価値、例えば満足度を算出する。質的評価規則は第1の実施の形態の同様の規則を用いる。

【0277】

すなわち、評価値算出機能46は、アクセスシーケンス・キーワードデータを読み込み、アクセスシーケンスごとに検索後閲覧時間を計算する。検索後閲覧時間は、同一のアクセスシーケンス識別子によるアクセスの中で、キーワード関連ページへのアクセスと一番新しいアクセスの時刻の差分から計算される。

【0278】

次に、評価値算出機能46は、上記表8のような検索後閲覧時間と満足度の対応データを読み込み、アクセスシーケンスごとの満足度を算出する。

【0279】

表23に、この段階における中間生成データを示す。

【0280】

【表23】

表23 中間生成データ

ViewID	VisitID	満足度	検索キーワード
View1	Visit1	0	
View2	Visit2	20	りんご
View3	Visit1	0	
View4	Visit2	20	
View5	Visit3	10	たまねぎ

次に、評価値算出機能46は、アクセスシーケンスについて満足度と検索キーワードを求める。この際、検索キーワードの記録のないアクセスシーケンスは処理対象としない。

【0281】

そして、評価値算出機能46は、検索キーワードごとに、満足度の平均を求め、関係データを生成する。

【0282】

関係データの具体的な内容は、上記表11と同様である。

【0283】

なお、本実施の形態において、評価値算出機能46は、量的評価値、例えばニーズ値を算出するとしてもよい。このニーズ値の算出方法は、上記第2の実施の形態で説明した方法が適用できる。また、評価値算出機能46によって質的評価値と量的評価値とを算出した場合、上記第3の実施の形態と同様にグラフ化機能27によって関係データに基づくグラフを表示するとしてもよい。

【0284】

以上説明した本実施の形態においては、アクセスシーケンス抽出、検索キーワード抽出、質的評価値算出及び量的評価値算出算出に関する各処理が並列ではなく直列に実行されており、上記各実施の形態と同様の効果を得ることができる。また、本実施の形態においては、ログを一括ではなく1アクセスシーケンスごとに処理することも可能になる。

【0285】

なお、上記各実施の形態においては、一つのアクセスシーケンスに対して検索

キーワードが一回入力された場合を想定して説明しているが、一つのアクセスシーケンスに対して検索キーワードが複数回入力された場合であっても検索キーワードの質的評価値、量的評価値を求めることができる。

【0286】

例えば、一つのアクセスシーケンスに対して検索キーワードが複数回入力されている場合、そのアクセスシーケンスを検索キーワードの入力を境界として複数のアクセスシーケンスに細分化し、この細分化されたアクセスシーケンスを用いて質的評価値、量的評価値を求めるとしてもよい。

【0287】

この場合、一つのアクセスシーケンスに対して検索キーワードが一回入力された場合における上記各実施の形態の処理方法を、細分化後の一つのアクセスシーケンスに対して適用することができる。

【0288】

また、上記各実施の形態に係る検索キーワード分析システムの各構成要素は、同様の作用を実現可能であれば配置を変更させてもよく、また各構成要素を自由に組み合わせてもよい。

【0289】

また、上記各実施の形態において、検索キーワード分析システムは複数の計算機により構成され、検索キーワード分析プログラムは複数の計算機に分散して配置され、互いに連携を取りつつ処理を実行するとしてもよい。

【0290】

本実施の形態に係る検索キーワード分析プログラムは、例えば磁気ディスク（フレキシブルディスク、ハードディスク等）、光ディスク（CD-ROM、DVD等）、半導体メモリなどの記録媒体に書き込んでコンピュータに適用可能である。また検索キーワード分析プログラムは、通信媒体により伝送してコンピュータに適用することも可能である。コンピュータは、記録媒体に記録された検索キーワード分析プログラムを読み込み、検索キーワード分析プログラムによって動作が制御されることにより、上記の機能を実現する。

【0291】

【発明の効果】

以上詳記したように本発明においては、閲覧者の検索結果に対する質的評価値を迅速かつ高精度に求めることができ、この質的評価値に基づいてWebページ又はWebサイトのデザインやコンテンツ、検索システムを改善することができる。

【図面の簡単な説明】**【図1】**

本発明の第1の実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図。

【図2】

同実施の形態に係る顧客サイトの構成の一例を示すブロック図。

【図3】

同実施の形態に係る検索キーワード分析システムの動作の一例を示すフローチャート。

【図4】

同実施の形態に係るキーワード抽出機能の動作の一例を示すフローチャート。

【図5】

同実施の形態に係る満足度算出機能の動作の一例を示すフローチャート。

【図6】

本発明の第2の実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図。

【図7】

本発明の第3の実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図。

【図8】

同実施の形態において表示されるグラフの一例を示す図。

【図9】

本発明の第4の実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図。

【図10】

本発明の第5の実施の形態に係る顧客サイトの構成の一例を示す図。

【図11】

サイト外検索システムから顧客サイトへの閲覧遷移の一例を示すブロック図。

【図12】

本発明の第6の実施の形態に係る顧客サイトの構成の一例を示す図。

【図13】

検索システムを用いた代表的なWebページの遷移の例を示す図。

【符号の説明】

4、38…記録装置

5…顧客

6…顧客サイト

6a…サイト内検索システム

7…Webアクセスログ

8…一般的な閲覧者

8a…顧客サイト閲覧者

9、18、24、29、42…検索キーワード分析システム

10、19、25、30、43…記録媒体

11、20、26、31、44…検索キーワード分析プログラム

12…ログ前処理機能

13…アクセスシーケンス抽出機能

14、45…キーワード抽出機能

15…満足度算出機能

16…満足度・キーワード関連付け機能

17、23、28、33…出力データ

21…ニーズ値算出機能

22…満足度・ニーズ値関連付け機能

27…グラフ化機能

32…成功度算出機能

341～34n…サイト外検索システム

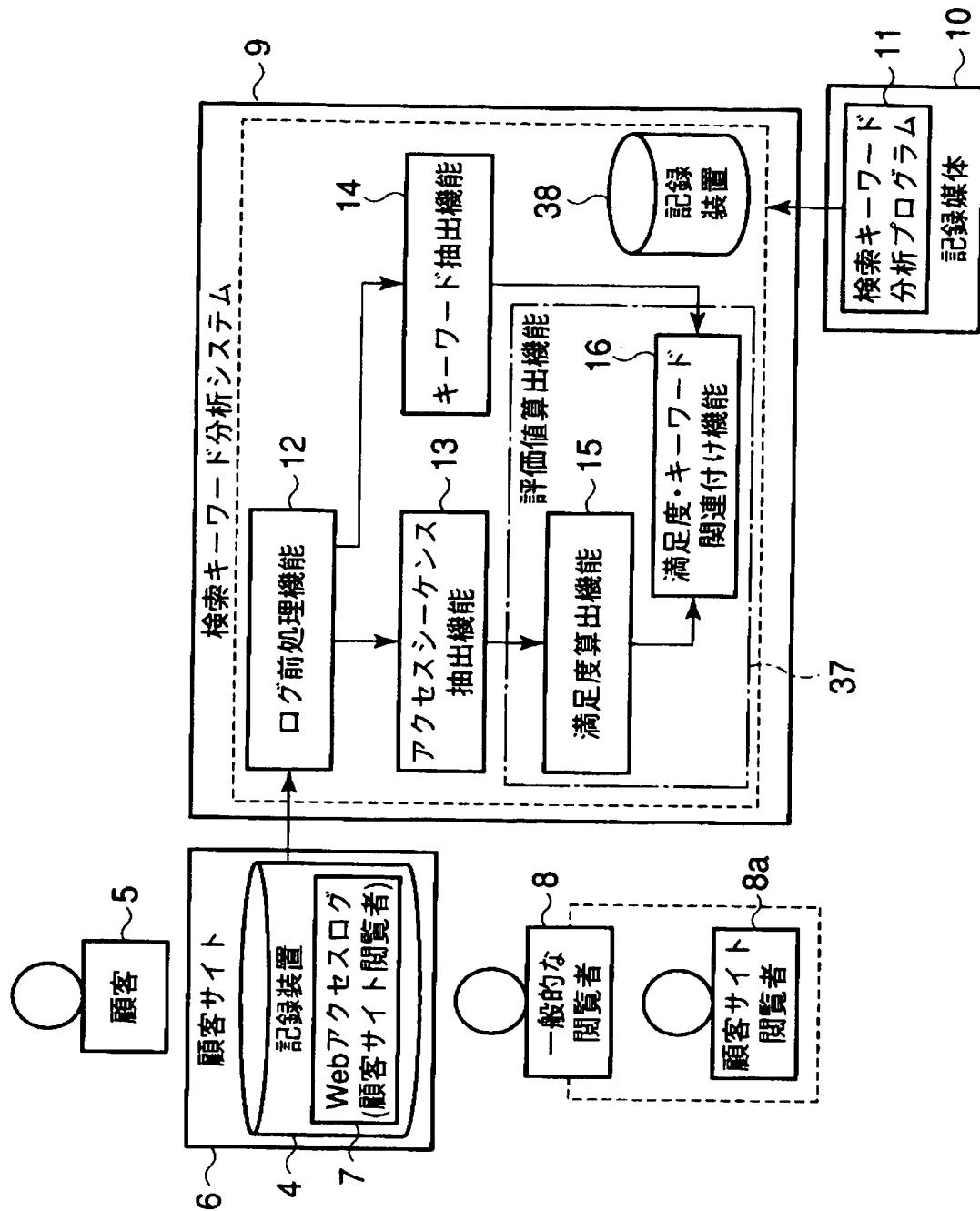
37、39、40、46…評価値算出機能

41…成功度・キーワード関連付け機能

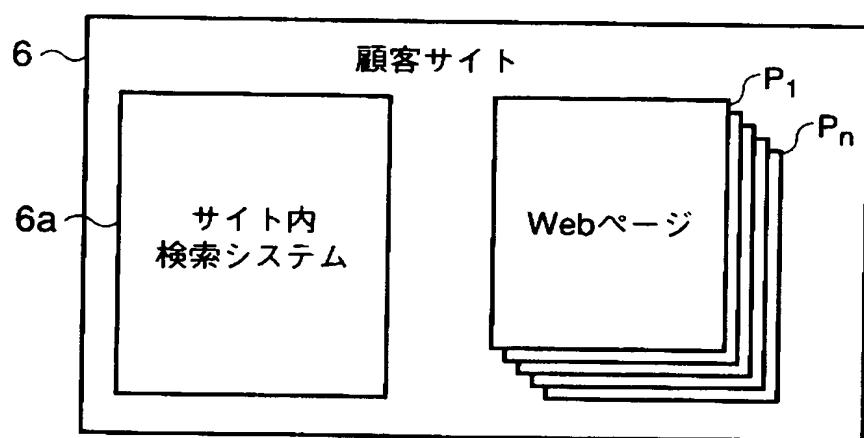
【書類名】

四面

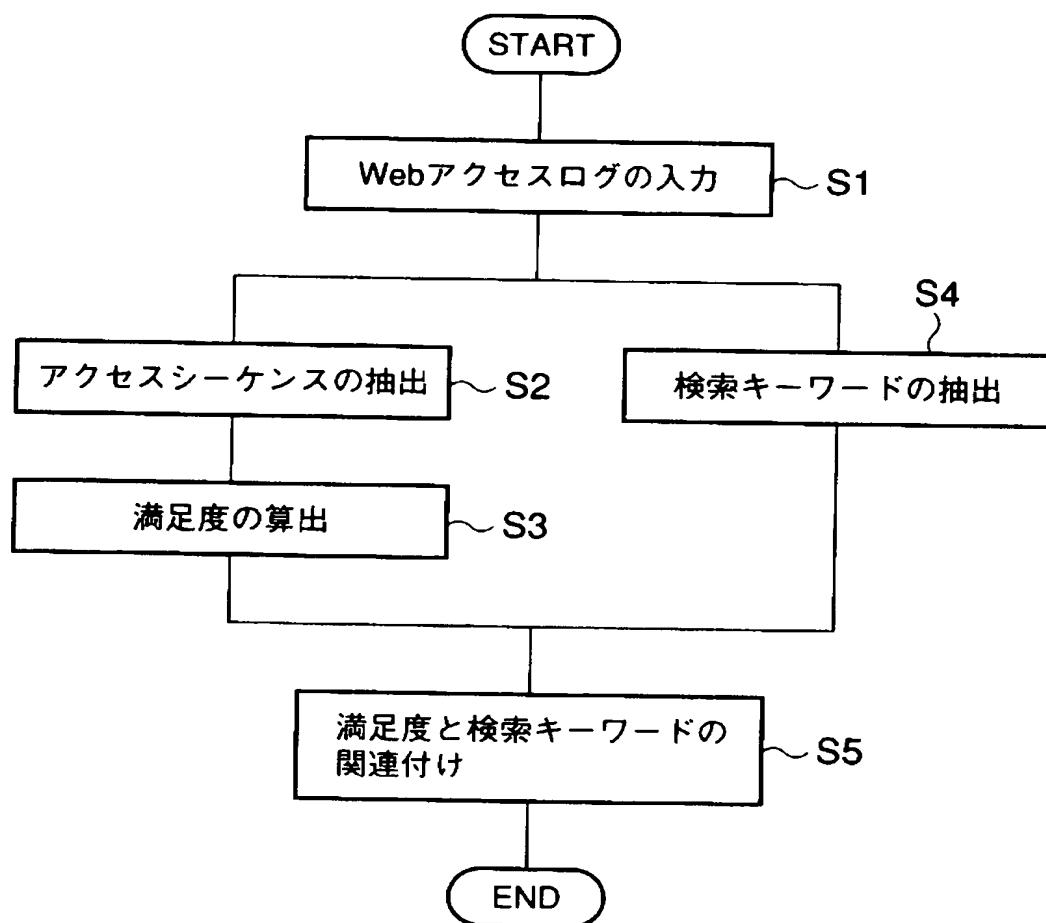
【図1】



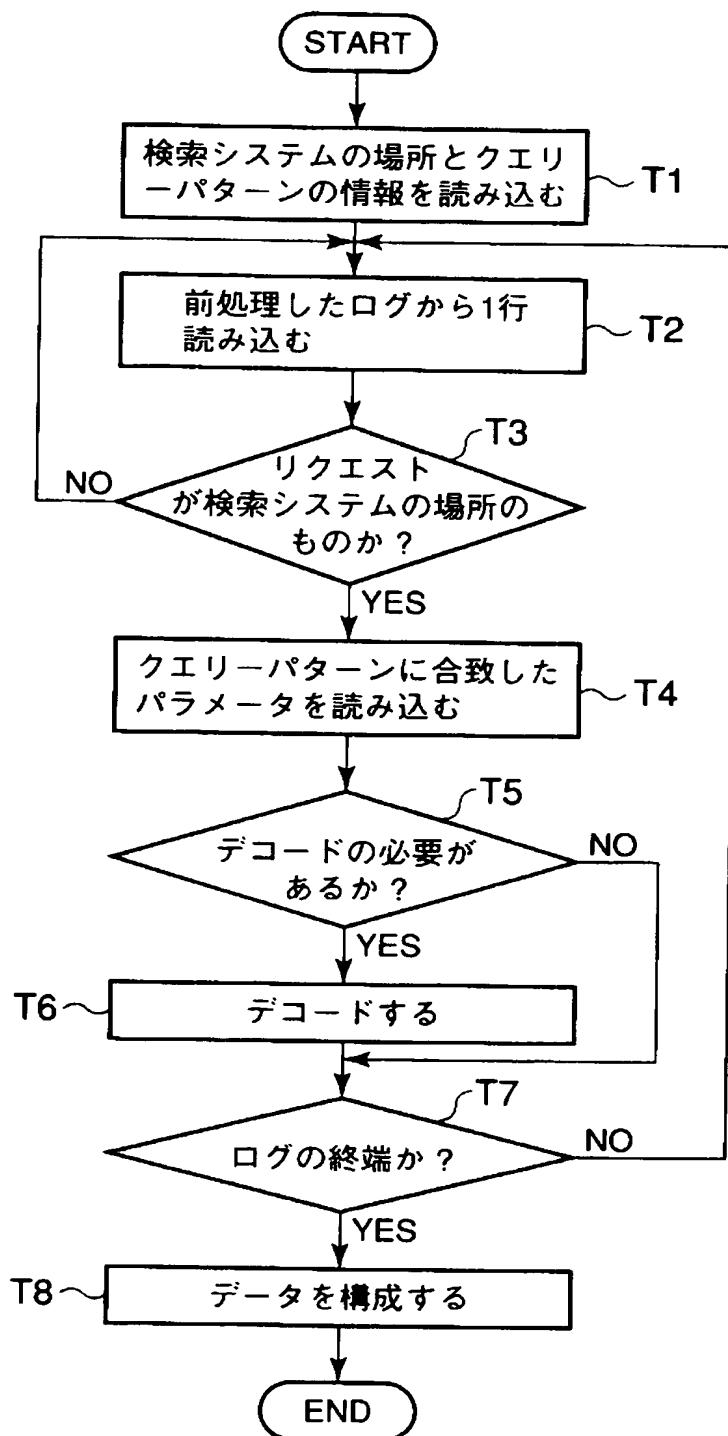
【図2】



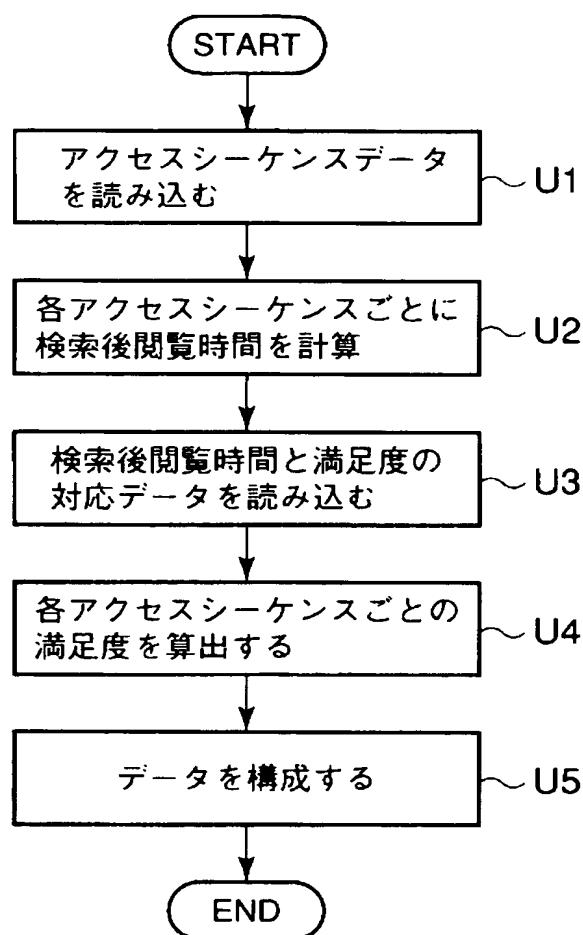
【図3】



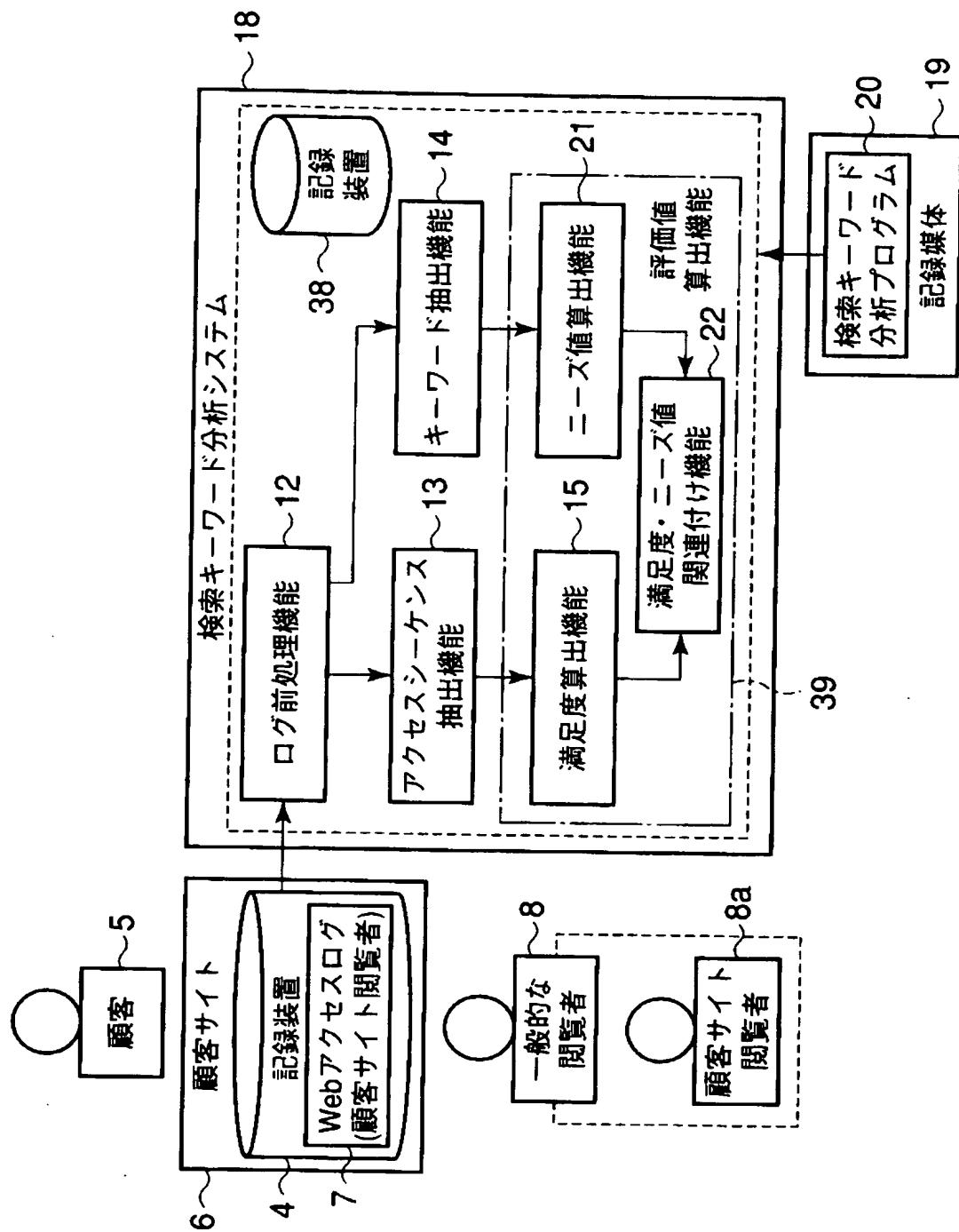
【図4】



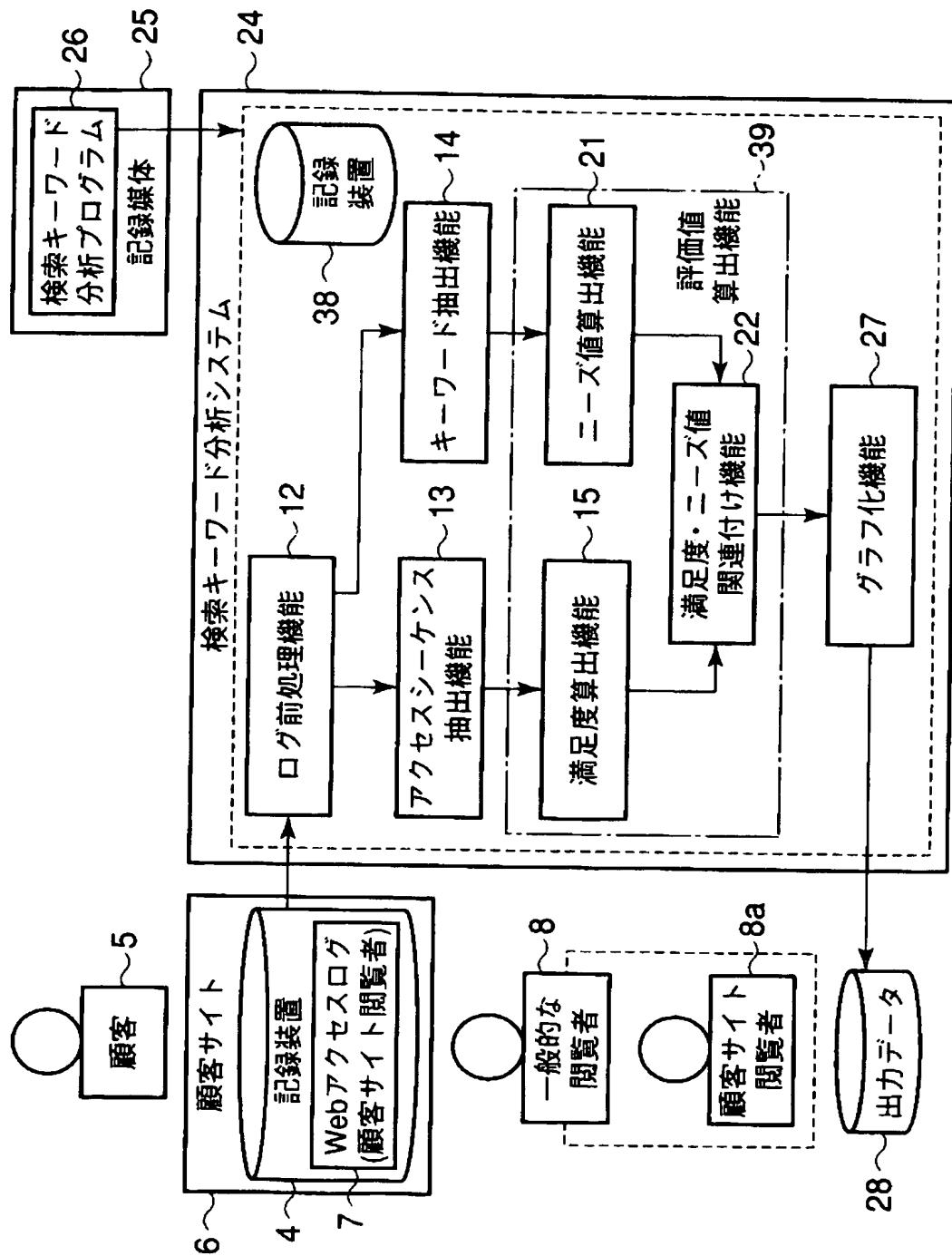
【図5】



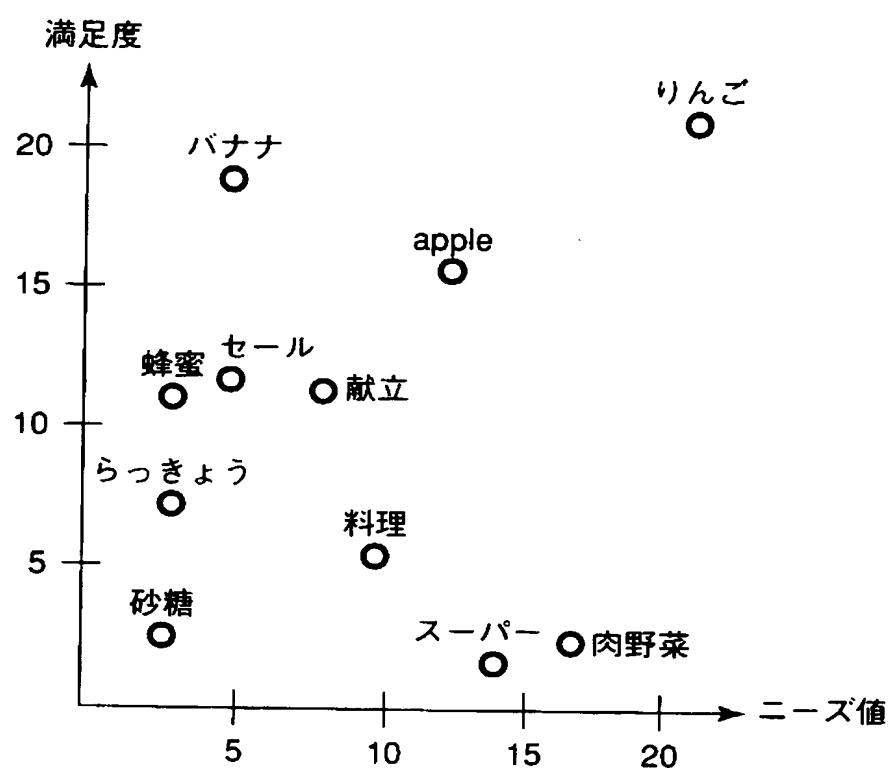
【図 6】



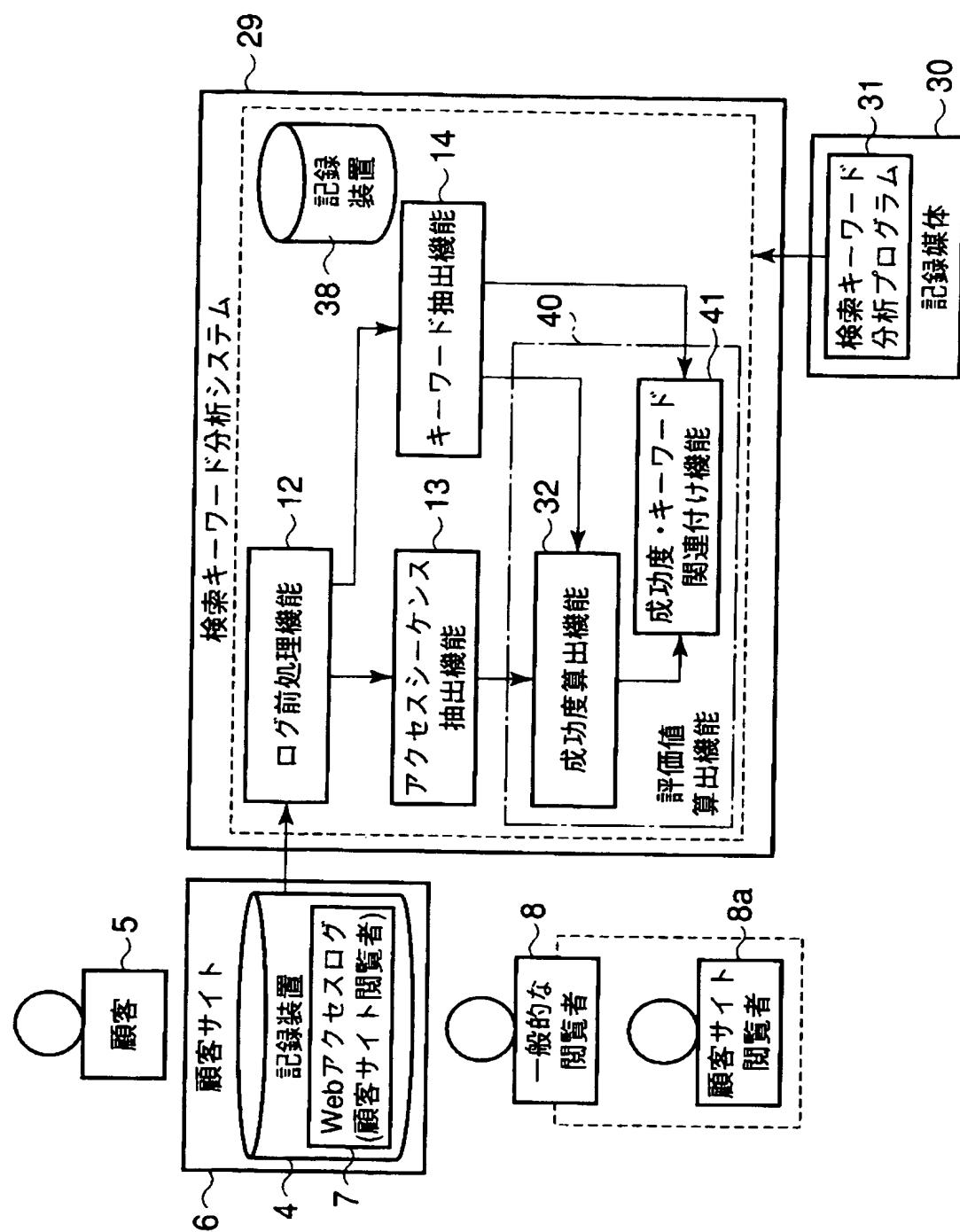
【図 7】



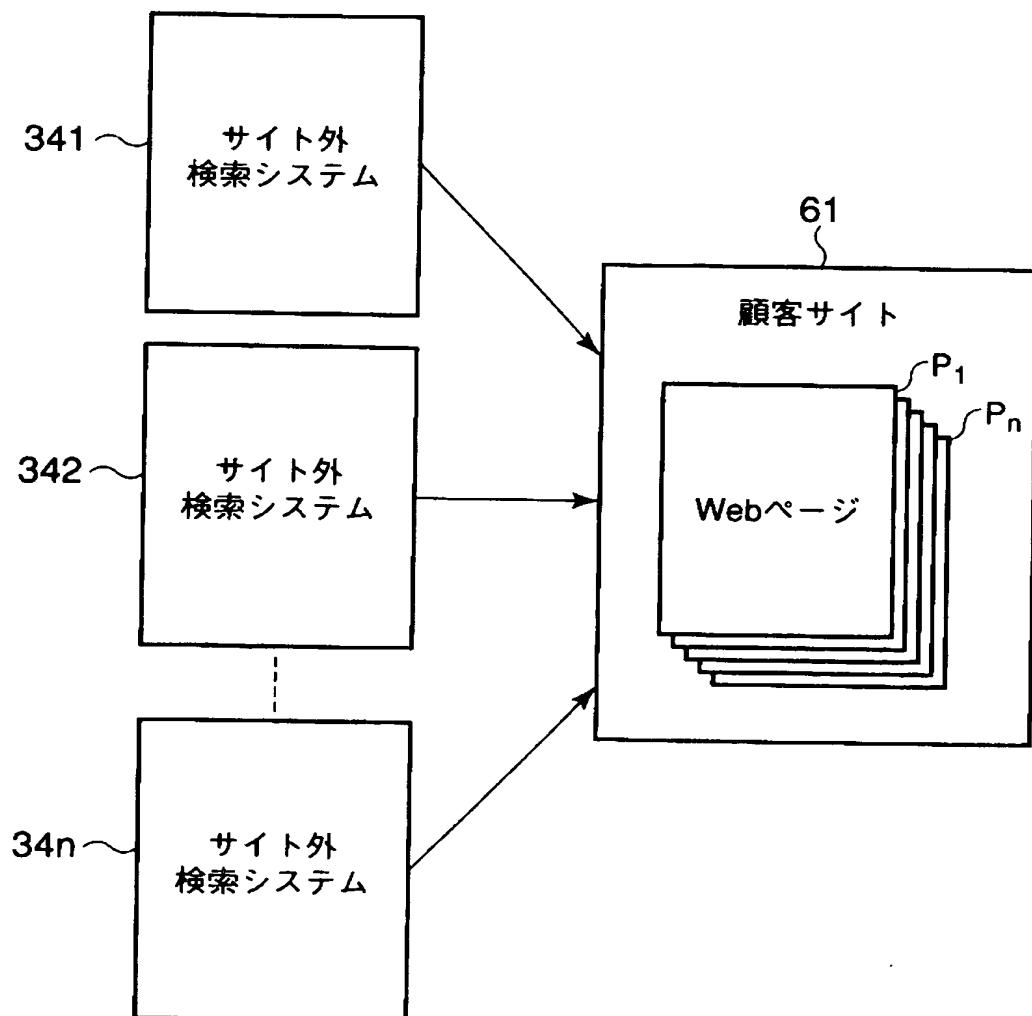
【図 8】



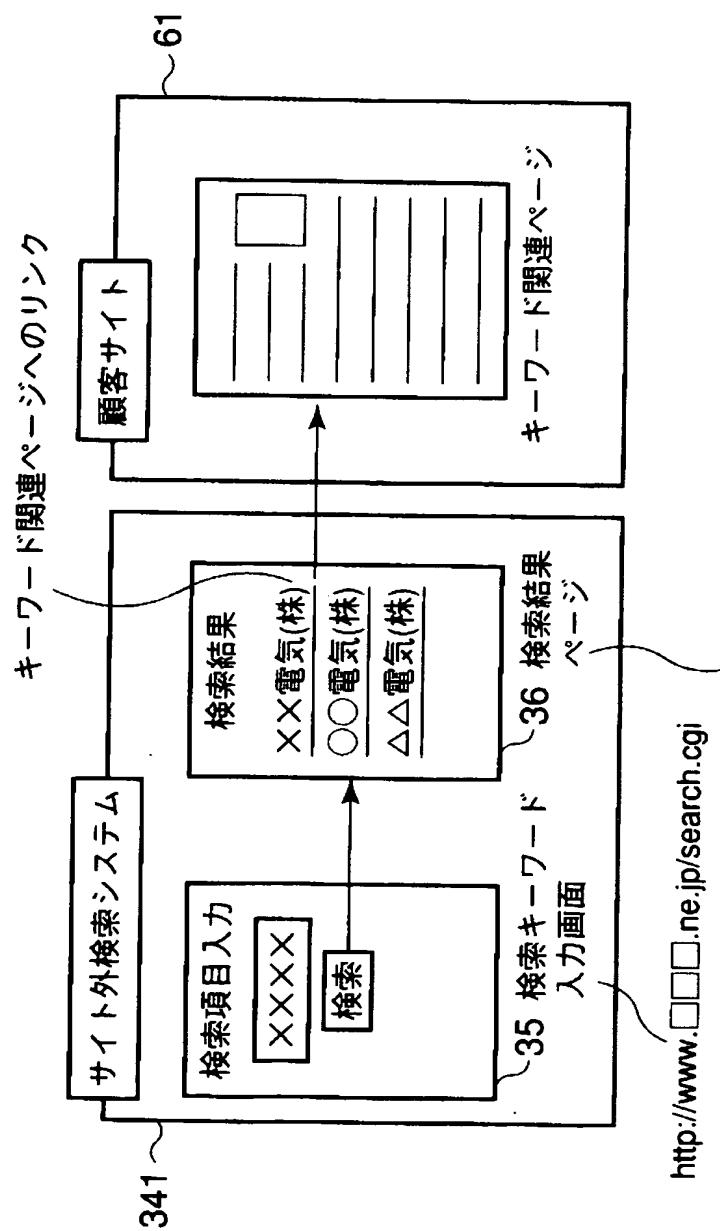
【図 9】



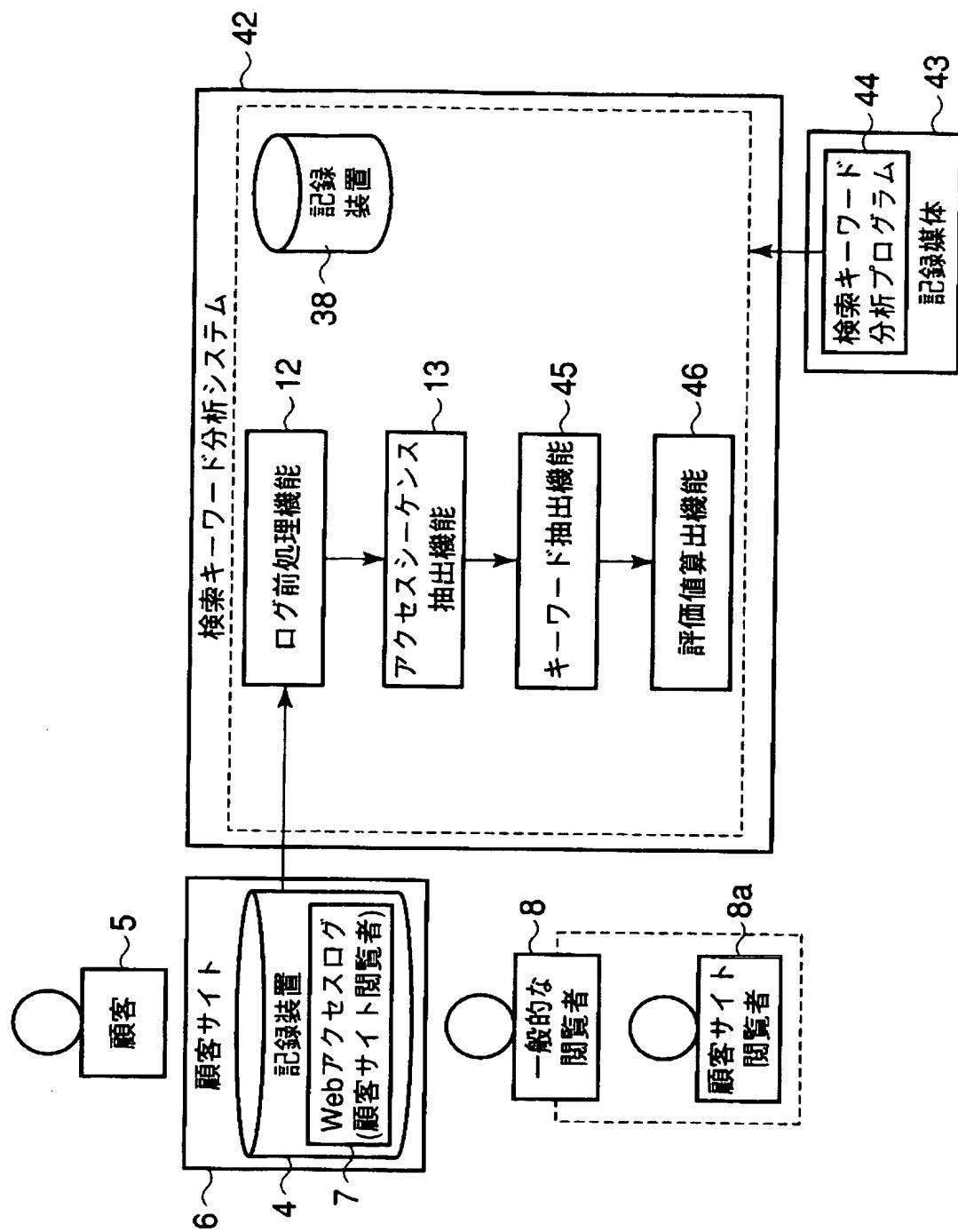
【図10】



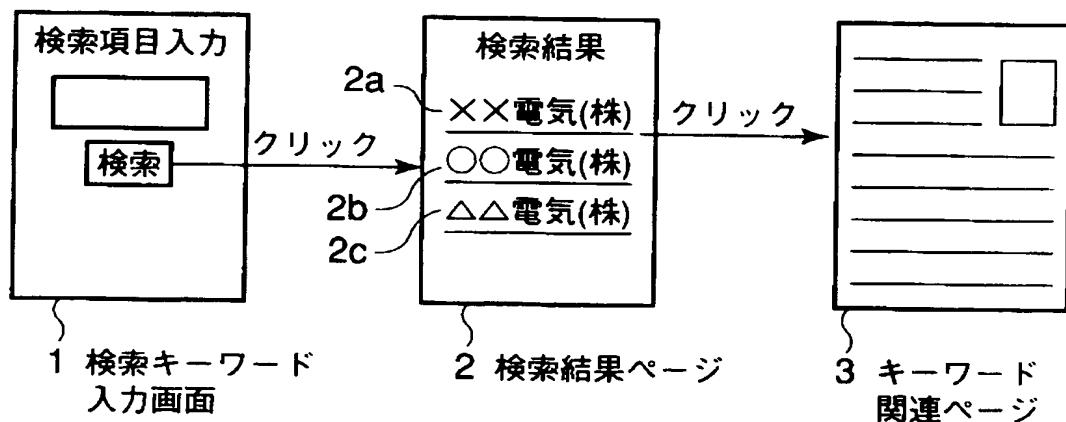
【図11】



【図12】



【図13】



【書類名】 要約書

【要約】

【課題】 入力された検索キーワードに基づく検索システムの検索結果に対する質的な評価値を求める。

【解決手段】 本発明のプログラム11は、コンピュータに、アクセス要求側を示すアクセス要求側識別データとアクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebログ7に基づいて同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出する機能13と、Webログ7において検索システムを示すアクセス先識別データと入力された検索キーワードとが関連付けられている場合に、検索キーワードを抽出する機能14と、アクセスの状態から検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいてアクセスシーケンスに対応する質的評価値を算出し、検索キーワードと質的評価値との対応関係を示す関係データを求める機能37とを実現させる。

【選択図】 図1

特願 2002-243466

出願人履歴情報

識別番号 [000003078]

1. 変更年月日 1990年 8月22日
[変更理由] 新規登録
住 所 神奈川県川崎市幸区堀川町72番地
氏 名 株式会社東芝

2. 変更年月日 2001年 7月 2日
[変更理由] 住所変更
住 所 東京都港区芝浦一丁目1番1号
氏 名 株式会社東芝

101645,623
0351021-1

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 2002年12月27日
Date of Application:

出願番号 特願2002-382396
Application Number:

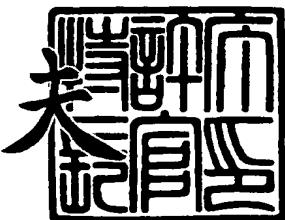
[ST. 10/C] : [JP2002-382396]

出願人 株式会社東芝
Applicant(s):

2003年9月2日

特許庁長官
Commissioner,
Japan Patent Office

今井康



出証番号 出証特2003-3071561

【書類名】 特許願
【整理番号】 A000206250
【提出日】 平成14年12月27日
【あて先】 特許庁長官 殿
【国際特許分類】 G06F 17/00
【発明の名称】 検索キーワード分析プログラム及びシステム並びに方法
【請求項の数】 30
【発明者】
【住所又は居所】 東京都府中市東芝町1番地 株式会社東芝府中事業所内
【氏名】 飯田 晴彦
【発明者】
【住所又は居所】 東京都府中市東芝町1番地 株式会社東芝府中事業所内
【氏名】 平井 潤
【発明者】
【住所又は居所】 東京都府中市東芝町1番地 株式会社東芝府中事業所内
【氏名】 平 博司
【特許出願人】
【識別番号】 000003078
【氏名又は名称】 株式会社 東芝
【代理人】
【識別番号】 100058479
【弁理士】
【氏名又は名称】 鈴江 武彦
【電話番号】 03-3502-3181
【選任した代理人】
【識別番号】 100084618
【弁理士】
【氏名又は名称】 村松 貞男

【選任した代理人】

【識別番号】 100068814

【弁理士】

【氏名又は名称】 坪井 淳

【選任した代理人】

【識別番号】 100092196

【弁理士】

【氏名又は名称】 橋本 良郎

【選任した代理人】

【識別番号】 100091351

【弁理士】

【氏名又は名称】 河野 哲

【選任した代理人】

【識別番号】 100088683

【弁理士】

【氏名又は名称】 中村 誠

【選任した代理人】

【識別番号】 100070437

【弁理士】

【氏名又は名称】 河井 将次

【先の出願に基づく優先権主張】

【出願番号】 特願2002-243466

【出願日】 平成14年 8月23日

【手数料の表示】

【予納台帳番号】 011567

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9705037

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 検索キーワード分析プログラム及びシステム並びに方法

【特許請求の範囲】

【請求項1】 コンピュータに、

WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、前記アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録するアクセスシーケンス抽出機能と、

前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けされている場合に、前記検索キーワードを抽出し、前記検索キーワードの抽出結果を示すキーワードデータを前記記録手段に記録するキーワード抽出機能と、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する質的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードのうち前記アクセスシーケンスにおいて入力された検索キーワードと前記質的評価値との対応関係を示す関係データを前記記録手段に記録する評価値算出機能と

を実現させるための検索キーワード分析プログラム。

【請求項2】 請求項1記載の検索キーワード分析プログラムにおいて、

前記評価値算出機能は、前記質的評価値に加えて、前記検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する量的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードのうち前記アクセスシーケンスにおいて入力された検索キーワード

ワードと前記質的評価値と前記量的評価値の対応関係を示す関係データを前記記録手段に記録する

ことを特徴とする検索キーワード分析プログラム。

【請求項3】 コンピュータに、

WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、前記アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録するアクセスシーケンス抽出機能と、

前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けられており前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスにおいて前記検索キーワードが入力されている場合に、前記検索キーワードを抽出し、前記検索キーワードの抽出結果を示すキーワードデータを前記記録手段に記録するキーワード抽出機能と、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対する質的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードと前記質的評価値との対応関係を示す関係データを前記記録手段に記録する評価値算出機能と

を実現させるための検索キーワード分析プログラム。

【請求項4】 請求項3記載の検索キーワード分析プログラムにおいて、

前記評価値算出機能は、前記質的評価値に加えて、前記検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対する量的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードと前記質的評価値と前記量的評価値の対応関係を示す関係データ

ータを前記記録手段に記録する
ことを特徴とする検索キーワード分析プログラム。

【請求項5】 請求項2又は請求項4記載の検索キーワード分析プログラムにおいて、

コンピュータに、前記記録手段に記録された関係データにしたがって、前記質的評価値と前記量的評価値に基づくグラフを描画するグラフ化機能を実現させるためのプログラムを附加したことを特徴とする検索キーワード分析プログラム。

【請求項6】 請求項2と請求項4と請求項5のうちのいずれか1項記載の検索キーワード分析プログラムにおいて、

前記量的評価値は、前記検索キーワードの入力数が多いほど値が大きくなる検索結果のニーズ値であり、

前記量的評価規則は、前記検索キーワードの入力数が多いほど前記ニーズ値が増加する規則であることを特徴とする検索キーワード分析プログラム。

【請求項7】 請求項1乃至請求項6のいずれか1項記載の検索キーワード分析プログラムにおいて、

前記Webアクセスログは、WWWにおけるアクセス毎に、前記アクセス要求側識別データと前記アクセス先識別データとアクセス時刻を示すデータとを関係付けており、

前記質的評価規則は、前記アクセスシーケンスにおける前記検索キーワード入力後のアクセスから最後のアクセスまでの時間間隔が長いほど前記アクセス要求側の満足度が増加する規則であり、

前記質的評価値は、前記満足度である
ことを特徴とする検索キーワード分析プログラム。

【請求項8】 請求項1乃至請求項6のいずれか1項記載の検索キーワード分析プログラムにおいて、

前記質的評価規則は、分析候補の検索キーワードと前記分析候補の検索キーワードの入力後にアクセスされた場合に成功度を増加するアクセス先識別データとを関連付けたデータと前記記録手段に記録されたキーワードデータとに基づいて

、前記アクセスシーケンスにおいて前記分析候補の検索キーワードが入力されておりその入力後前記分析候補の検索キーワードと関連付けされているアクセス先識別データの示すアクセス先がアクセスされている場合に、前記アクセスシーケンスに対する成功度を増加する規則であり、

前記質的評価値は、前記成功度であることを特徴とする検索キーワード分析プログラム。

【請求項9】 請求項8に記載の検索キーワード分析プログラムにおいて、前記質的評価規則は、前記分析候補の検索キーワードが入力されてから前記分析候補の検索キーワードと関連付けされているアクセス先識別データの示すアクセス先がアクセスされるまでのアクセス数が少ないほど加えられる成功度を大きくすることを特徴とする検索キーワード分析プログラム。

【請求項10】 請求項1乃至請求項9のいずれか1項記載の検索キーワード分析プログラムにおいて、

前記Webアクセスログは、前記アクセス先識別データの示すアクセス先がハイパーリンクを使ってアクセスされた場合に、前記アクセス先識別データの示すアクセス先の前段でアクセスされたアクセス先を示すレファラ情報と前記アクセス先識別データとを関連付けたデータを含み、

前記アクセスシーケンス抽出機能は、前記レファラ情報が前記検索システムを示す場合に、前記レファラ情報に関連付けされている前記アクセス先識別データの示すアクセス先へのアクセスを含むアクセスシーケンスを抽出し、

前記キーワード抽出機能は、前記レファラ情報が前記検索システムを示す場合に、前記レファラ情報に含まれている検索キーワードを抽出し、

前記評価値算出機能は、前記アクセスシーケンスに対する前記検索キーワード入力後の質的評価値を算出する

ことを特徴とする検索キーワード分析プログラム。

【請求項11】 請求項2又は請求項4記載の検索キーワード分析プログラムにおいて、

前記Webアクセスログは、前記アクセス要求側識別データと前記アクセス先識別データとアクセスされた時間を示す時間データとを関係付けたデータであり

前記評価値算出機能は、前記時間データに基づいて、複数の期間毎の前記関係データを求め、

コンピュータに、

前記複数の期間毎の関係データに基づいて、前記検索キーワードについて前記質的評価値と前記量的評価値との遷移状況を表すトレンドデータを求め、前記記録手段に記録するトレンド分析機能

を実現させることを特徴とする検索キーワード分析プログラム。

【請求項 1 2】 請求項 1 1 記載の検索キーワード分析プログラムにおいて

前記トレンド分析機能は、前記複数の期間の切り換わりに応じて、前記検索キーワードの前記質的評価値と前記量的評価値とのうち少なくとも一方が、前記質的評価値と前記量的評価値のそれぞれについて設定された境界値を超えて変化するか判断し、

前記トレンドデータは、前記トレンド分析機能による判断結果を示すデータを含む

ことを特徴とする検索キーワード分析プログラム。

【請求項 1 3】 請求項 1 2 記載の検索キーワード分析プログラムにおいて

前記トレンド分析機能は、

前記複数の期間毎に、前記複数の境界値から定まる複数の領域のうち、前記検索キーワードの属する領域を求める領域計算機能と、

前記複数の期間の切り換わりに応じて前記検索キーワードの領域が変化するか判断する領域判断機能と、

前記領域が変化する検索キーワードに対して前記領域の変化に応じたメッセージを付す領域メッセージ付加機能と
を具備し、

前記トレンドデータは、前記領域が変化する検索キーワードに前記メッセージを付した結果を示すデータを含む

ことを特徴とする検索キーワード分析プログラム。

【請求項14】 請求項11乃至請求項13のいずれか1項記載の検索キーワード分析プログラムにおいて、

前記トレンド分析機能は、前記複数の期間毎の前記検索キーワードの前記質的評価値と前記量的評価値とに基づいて、前記複数の期間の切り換わり前後における前記検索キーワードの移動距離と移動方向とを求め、

前記トレンドデータは、前記複数の期間の切り換わり前後における前記検索キーワードの移動距離と移動方向を示すデータを含むことを特徴とする検索キーワード分析プログラム。

【請求項15】 請求項11乃至請求項14のいずれか1項記載の検索キーワード分析プログラムにおいて、

前記トレンド分析機能は、

前記複数の期間毎の前記検索キーワードの前記質的評価値と前記量的評価値とに基づいて、前記複数の期間の切り換わり前後における前記検索キーワードの移動距離を求める移動距離計算機能と、

前記移動距離が移動距離しきい値を超えているか判断する移動距離判断機能と

前記移動距離が前記移動距離しきい値を超えている場合に、前記検索キーワードの移動方向を求める移動方向計算機能と、

前記移動距離が前記移動距離しきい値を超える検索キーワードに対して前記移動方向に応じたメッセージを付すべくトルメッセージ付加機能とを具備し、

前記トレンドデータは、前記移動距離が前記移動距離しきい値を超える検索キーワードに前記メッセージを付した結果を示すデータを含むことを特徴とする検索キーワード分析プログラム。

【請求項16】 請求項11乃至請求項15のいずれか1項記載の検索キーワード分析プログラムにおいて、

コンピュータに、

前記トレンドデータに基づいて、前記質的評価値と前記量的評価値とを複数の

座標軸とする空間上に、前記複数の期間毎の前記検索キーワードを示す指標を前記質的評価値と前記量的評価値とに応じて配置し、配置された前記指標を前記複数の期間の経過順序に応じて線で結んだグラフを表示するための処理を実行する出力処理機能を実現させる

ことを特徴とする検索キーワード分析プログラム。

【請求項 17】 請求項2又は請求項4記載の検索キーワード分析プログラムにおいて、

前記Webアクセスログは、前記アクセス要求側識別データと前記アクセス先識別データとアクセスされた時間を示す時間データとを関係付けたデータであり

前記評価値算出機能は、前記時間データに基づいて、複数の期間毎の前記関係データを求め、

コンピュータに、

前記検索キーワードと前記検索キーワードの属するカテゴリとを関係付けたカテゴリデータと前記複数の期間毎の関係データとにに基づいて、前記複数の期間毎に前記カテゴリについて前記質的評価値と前記量的評価値とを算出し、前記記録手段に記録するカテゴリ形成機能と、

前記カテゴリについて前記質的評価値と前記量的評価値との遷移状況を表すトレンドデータを求め、前記記録手段に記録するトレンド分析機能とを実現させることを特徴とする検索キーワード分析プログラム。

【請求項 18】 請求項17記載の検索キーワード分析プログラムにおいて

コンピュータに、

前記トレンドデータに基づいて、前記質的評価値と前記量的評価値とを複数の座標軸とする空間上に、前記複数の期間毎の前記カテゴリを示す指標を前記質的評価値と前記量的評価値とに応じて配置し、配置された前記指標を前記複数の期間の経過順序に応じて線で結んだグラフを表示するための処理を実行する出力処理機能を実現させる

ことを特徴とする検索キーワード分析プログラム。

【請求項 19】 WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出するアクセスシーケンス抽出手段と、

前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けされている場合に、前記検索キーワードを抽出するキーワード抽出手段と、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記アクセスシーケンスに対応する質的評価値を算出し、前記アクセスシーケンスにおいて入力された検索キーワードと前記質的評価値との対応関係を示す関係データを作成する評価値算出手段と

を具備する検索キーワード分析システム。

【請求項 20】 請求項 19 記載の検索キーワード分析システムにおいて、前記評価値算出手段は、前記質的評価値に加えて、前記検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、前記アクセスシーケンスに対応する量的評価値を算出し、前記アクセスシーケンスにおいて入力された検索キーワードと前記質的評価値と前記量的評価値の対応関係を示す関係データを作成する

ことを特徴とする検索キーワード分析システム。

【請求項 21】 WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出するアクセスシーケンス抽出手段と、

前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワード

とが関連付けされており前記アクセスシーケンスにおいて前記検索キーワードが入力されている場合に、前記検索キーワードを抽出するキーワード抽出手段と、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記アクセスシーケンスに対応する質的評価値を算出し、前記検索キーワードと前記質的評価値との対応関係を示す関係データを作成する評価値算出手段と
を具備する検索キーワード分析システム。

【請求項22】 請求項21記載の検索キーワード分析システムにおいて、

前記評価値算出手段は、前記質的評価値に加えて、前記検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、前記アクセスシーケンスに対応する量的評価値を算出し、前記検索キーワードと前記質的評価値と前記量的評価値の対応関係を示す関係データを作成することを特徴とする検索キーワード分析システム。

【請求項23】 請求項20又は請求項22記載の検索キーワード分析システムにおいて、

前記Webアクセスログは、前記アクセス要求側識別データと前記アクセス先識別データとアクセスされた時間を示す時間データとを関係付けたデータであり

前記評価値算出手段は、前記時間データに基づいて、複数の期間毎の前記関係データを求め、

前記複数の期間毎の関係データに基づいて、前記検索キーワードについて前記質的評価値と前記量的評価値との遷移状況を表すトレンドデータを求めるトレンド分析手段を具備することを特徴とする検索キーワード分析システム。

【請求項24】 請求項20又は請求項22記載の検索キーワード分析システムにおいて、

前記Webアクセスログは、前記アクセス要求側識別データと前記アクセス先識別データとアクセスされた時間を示す時間データとを関係付けたデータであり

前記評価値算出手段は、前記時間データに基づいて、複数の期間毎の前記関係

データを求め、

前記検索キーワードと前記検索キーワードの属するカテゴリとを関係付けたカテゴリデータと前記複数の期間毎の関係データとに基づいて、前記複数の期間毎に前記カテゴリについて前記質的評価値と前記量的評価値とを算出し、前記記録手段に記録するカテゴリ形成手段と、

前記カテゴリについて前記質的評価値と前記量的評価値との遷移状況を表すトレンドデータを求めるトレンド分析手段と
を具備する検索キーワード分析システム。

【請求項 25】 WWWにおける検索システムに入力された検索キーワードについて分析するためにコンピュータにより実施される検索キーワード分析方法において、

WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、前記アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録するとともに、前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けされている場合に、前記検索キーワードを抽出し、前記検索キーワードの抽出結果を示すキーワードデータを前記記録手段に記録し、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する質的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードのうち前記アクセスシーケンスにおいて入力された検索キーワードと前記質的評価値との対応関係を示す関係データを前記記録手段に記録することを特徴とする検索キーワード分析方法。

【請求項 26】 請求項 25 記載の検索キーワード分析方法において、
前記質的評価値とともに、前記検索システムの検索結果に対する量的評価値を

算出する基準を定義した量的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する量的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードのうち前記アクセスシーケンスにおいて入力された検索キーワードと前記質的評価値と前記量的評価値の対応関係を示す関係データを前記記録手段に記録することを特徴とする検索キーワード分析方法。

【請求項27】 WWWにおける検索システムに入力された検索キーワードについて分析するためにコンピュータにより実施される検索キーワード分析方法において、

WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、前記アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録し、

前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けられており前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスにおいて前記検索キーワードが入力されている場合に、前記検索キーワードを抽出し、前記検索キーワードの抽出結果を示すキーワードデータを前記記録手段に記録し、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する質的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードと前記質的評価値との対応関係を示す関係データを前記記録手段に記録することを特徴とする検索キーワード分析方法。

【請求項28】 請求項27記載の検索キーワード分析方法において、

前記質的評価値に加えて、前記検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、前記記録手段に記録されたア

クセスシーケンスデータの示すアクセスシーケンスに対応する量的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードと前記質的評価値と前記量的評価値の対応関係を示す関係データを前記記録手段に記録する

ことを特徴とする検索キーワード分析方法。

【請求項29】 請求項26又は請求項28記載の検索キーワード分析方法において、

前記Webアクセスログは、前記アクセス要求側識別データと前記アクセス先識別データとアクセスされた時間を示す時間データとを関係付けたデータであり

前記関係データは、前記時間データに基づいて、複数の期間毎に求められ、前記複数の期間毎の関係データに基づいて、前記検索キーワードについて前記質的評価値と前記量的評価値との遷移状況を表すトレンドデータを求め、前記記録手段に記録する

ことを特徴とする検索キーワード分析方法。

【請求項30】 請求項26又は請求項28記載の検索キーワード分析方法において、

前記Webアクセスログは、前記アクセス要求側識別データと前記アクセス先識別データとアクセスされた時間を示す時間データとを関係付けたデータであり

前記関係データは、前記時間データに基づいて、複数の期間毎に求められ、前記検索キーワードと前記検索キーワードの属するカテゴリとを関係付けたカテゴリデータと前記複数の期間毎の関係データとにに基づいて、前記複数の期間毎に前記カテゴリについて前記質的評価値と前記量的評価値とを算出し、前記記録手段に記録し、

前記カテゴリについて前記質的評価値と前記量的評価値との遷移状況を表すトレンドデータを求め、前記記録手段に記録する

ことを特徴とする検索キーワード分析方法。

【発明の詳細な説明】

【0001】**【発明の属する技術分野】**

本発明は、WWW (World Wide Web) において、目的のWebページ又はWebサイトを探すために、Web閲覧者（以下、単に「閲覧者」という）からWWW検索システム（以下、「検索システム」という）に入力された検索キーワードの特徴を分析するためのプログラム及びシステム並びに方法に関する。

【0002】**【従来の技術】**

インターネットにおけるアクセス先の例として、Webページ又はWebサイトがある。

【0003】

インターネット上のWebページ又はWebサイトは、宣伝広告、情報発信、受注受付などの様々な目的のために運営される。かかる目的を効果的に達成するために、Webページ又はWebサイトが分析、調査される。

【0004】

例えば、分析者は、分析対象のWebサイトに対し、どのような他のWebサイトからアクセスされたか知ることにより、バナー広告、ディレクトリサイトへの掲載、懸賞サイトへの出品などの戦略を決定できる。

【0005】

また、分析者は、検索エンジンでどのような検索が実行され、分析対象のWebサイトがアクセスされたかなどを知ることにより、コンテンツ改善の指針を決定できる。

【0006】

Webページ又はWebサイトの分析、調査の方法として、どのWebページが、いつ、どのようにアクセスされたかを記録するWebアクセスログを用いたWebアクセスログ分析がある。

【0007】

Webアクセスログには、主に2種類ある。

【0008】

第1のWebアクセスログは、单一又は複数のWebサーバに記録されたWebアクセスログ（狭義のWebアクセスログ）である。

【0009】

第2のWebアクセスログは、一人又は複数の閲覧者による様々なWebページ又はWebサイトへのアクセスを記録したWebアクセスログである。この第2のWebアクセスログには、例えば、WWW視聴率調査会社がパネル（モニター）から収集した、各パネルによるWebページ又はWebサイトへのアクセスログがある。

【0010】

以下において、Webアクセスログとは、第1のWebアクセスログを意味するとして説明する。

【0011】

例えば、Webアクセスログ分析では、Webページ又はWebサイトにアクセスした閲覧者数（ユーザ数、訪問者数）が、時間毎、曜日毎にどのように変化するか調査される。

【0012】

また、Webアクセスログ分析では、閲覧者が使用しているブラウザ又はコンピュータのOSの種類が調査される場合もある。

【0013】

また、Webアクセスログ分析では、一つのWebサイト内又は複数のWebサイト間をまたがって閲覧者がどのような閲覧行動をとるかについて調査される場合もある。例えば、閲覧者がWebページAを閲覧し、次にWebページBを閲覧したことを明らかにするなどのような経路分析もWebアクセスログ分析に含まれる。

【0014】

一方、インターネット上のWebページ又はWebサイトの閲覧を補助するために、閲覧者の入力した検索キーワード（例えば、「車」「ペット」など）に応じて、この検索キーワードに関係する内容のWebページ（以下、「キーワード関連ページ」という）又はWebサイト（以下、「キーワード関連サイト」とい

う)についての情報（例えば、キーワード関連ページ又はキーワード関連サイトのURL（Uniform Resource Locators）、要約情報、リンク情報など）のリストを検索結果ページとして表示する検索システムが存在する。この検索システムは、広く利用されている。

【0015】

図31は、検索システムを用いた代表的なWebページの遷移の例を示す。

【0016】

検索キーワード入力画面1上で閲覧者のブラウザから検索キーワードが入力され、検索開始が指示されると、検索システムは、キーワード関連ページへのリンク2a～2cを抽出し、抽出されたキーワード関連ページへのリンク2a～2cを表示する検索結果ページ2を閲覧者に対して表示する。

【0017】

検索結果ページ2上のいずれかのリンクがクリックされると、このクリックされたリンクの先のキーワード関連ページが閲覧者のブラウザに表示される。例えば、リンク2aがクリックされると、キーワード関連ページ3が表示される。

【0018】

この検索システムを用いた場合、WebサーバのWebアクセスログには、検索キーワードとその入力日時などを含む付加情報が記録される。

【0019】

したがって、このWebアクセスログを用いて、検索キーワードに関する分析（以下、「検索キーワード分析」という）が可能となる。

【0020】

検索キーワード分析においては、例えば、ある検索キーワードが何回入力されたか調査される。

【0021】

また、例えば、検索結果ページにおいて検索結果が実際にクリックされたか調査される。

【0022】

また、例えば、ある検索キーワードに対して、キーワード関連ページが何件抽

出されたか調査される。

【0023】

特開2002-123517号（特許文献1）には、検索キーワード分析技術、商用ツール、非商用ツール、サービスが記載されている。

【0024】

【特許文献1】

特開2002-123517号公報

【0025】

【発明が解決しようとする課題】

上述したように、Webアクセスログ分析では、レファラ情報などを用いて分析対象のWebサイトをアクセスする前のWebサイトの分析、検索キーワード単位での閲覧者数の分析などは行なわれている。

【0026】

しかしながら、従来のWebアクセスログ分析では、分析者は、分析対象のWebサイトに閲覧者が満足しているか、分析対象のWebサイトにとっての成功（目的）がどの程度達成されているかまで理解することが困難である。

【0027】

例えば、分析対象のWebサイトは商品の販売数向上を目的としているとする。従来のWebアクセスログでは、検索エンジンの検索結果に基づいて分析対象のWebサイトがアクセスされた場合にこの検索エンジンに入力された検索キーワード毎の閲覧者数を求めることはできる。

【0028】

しかし、閲覧者が多くても分析対象のWebサイトで全く商品が購入されない場合もあり、閲覧者が少なくてもほとんどの閲覧者が商品を購入する場合もある。そのため、分析者は、従来のWebアクセスログ分析の結果から、その検索キーワードを入力した閲覧者が分析対象のWebサイトに満足したか、どの検索キーワードを入力した閲覧者に対してWebサイトの目的が達成されたか判断することが困難である。

【0029】

また、例えば、従来のWebアクセスログ分析により、ある検索キーワードに對して、キーワード関連ページが何件抽出されたか調査することができる。そして、分析者は、閲覧者の希望するWebページの抽出漏れを防止する観点から、キーワード関連ページの抽出数が多いほど検索結果に対する閲覧者の満足度が高く、少ないほど満足度が低いと判断する場合がある。

【0030】

しかしながら、キーワード関連ページの抽出数は、閲覧者の満足度などを直接表すわけではない。実際には、キーワード関連ページの抽出数が少ない方が閲覧者にとってはリンクを選択しやすく、閲覧者の希望に合致し、満足度が高い場合もある。すなわち、閲覧者の希望と合致するキーワード関連ページが検索結果として閲覧者に提示される場合、閲覧者は、冗長な検索結果が提示されるよりもキーワード関連ページの抽出数が少ない程、検索結果に対して満足する。

【0031】

したがって、分析者は、従来のWebアクセスログ分析から閲覧者が検索結果に満足しているか判断することが困難である。

【0032】

また、例えば、検索結果ページ上のキーワード関連ページへのリンクが実際にクリックされたか、クリックされている場合にはどのキーワード関連ページへのリンクがクリックされたか調査すると、閲覧者によって検索結果ページ上でどのキーワード関連ページが選択されたか知ることができる。

【0033】

しかしながら、従来のWebアクセスログ分析により、アクセスされたキーワード関連ページが本当に閲覧者の求めた内容を掲載しており閲覧者が満足したかについて、分析者が判断すること（検索結果に対して質的な評価をすること）は極めて困難である。

【0034】

本発明は、以上のような実情に鑑みてなされたもので、Webアクセスログ分析において分析者の求める分析結果を適切に提供する検索キーワード分析プログラム及びシステム並びに方法を提供することを目的とする。

【0035】

【課題を解決するための手段】

本発明を実現するにあたって講じた具体的手段について以下に説明する。

【0036】

第1の発明の検索キーワード分析プログラムは、コンピュータに、WWWにおけるアクセス要求側を示すアクセス要求側識別データとアクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録するアクセスシーケンス抽出機能と、WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データとアクセス要求側から検索システムに入力された検索キーワードとが関連付けられている場合に、検索キーワードを抽出し、検索キーワードの抽出結果を示すキーワードデータを記録手段に記録するキーワード抽出機能と、WWWにおけるアクセスの状態から検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する質的評価値を算出し、記録手段に記録されているキーワードデータの示す検索キーワードのうちアクセスシーケンスにおいて入力された検索キーワードと質的評価値との対応関係を示す関係データを記録手段に記録する評価値算出機能とを実現させる。

【0037】

これにより、入力された検索キーワードに基づく検索システムの検索結果に対する質的評価値を容易に把握できる。この第1の発明では、アクセスシーケンス抽出機能とキーワード抽出機能とを並列に動作させることができる。

【0038】

なお、第1の発明において、評価値算出機能は、質的評価値に加えて、検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する量的評価値を算出し、記録手段に記録されているキーワードデ

ータの示す検索キーワードのうちアクセスシーケンスにおいて入力された検索キーワードと質的評価値と量的評価値の対応関係を示す関係データを記録手段に記録するとしてもよい。

【0039】

これにより、入力された検索キーワードに基づく検索システムの検索結果に対する質的評価値、量的評価値、質的評価値と量的評価値との関係を容易に把握できる。

【0040】

第2の発明の検索キーワード分析プログラムは、コンピュータに、WWWにおけるアクセス要求側を示すアクセス要求側識別データとアクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録するアクセスシーケンス抽出機能と、WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データとアクセス要求側から検索システムに入力された検索キーワードとが関連付けられており記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスにおいて検索キーワードが入力されている場合に、検索キーワードを抽出し、検索キーワードの抽出結果を示すキーワードデータを記録手段に記録するキーワード抽出機能と、WWWにおけるアクセスの状態から検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する質的評価値を算出し、記録手段に記録されているキーワードデータの示す検索キーワードと質的評価値との対応関係を示す関係データを記録手段に記録する評価値算出機能とを実現させる。

【0041】

この第2の発明では、上記第1の発明と同様の効果を得ることができるが、アクセスシーケンス抽出機能とキーワード抽出機能とが直列に動作する特徴を持つ。

【0042】

なお、第2の発明において、評価値算出機能は、質的評価値に加えて、検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対する量的評価値を算出し、記録手段に記録されているキーワードデータの示す検索キーワードと質的評価値と量的評価値の対応関係を示す関係データを前記記録手段に記録するとしてもよい。

【0043】

上記第1又は第2の発明の検索キーワード分析プログラムにおいて、検索キーワードと質的評価値と量的評価値の対応関係を示す関係データが求められる場合、コンピュータに、記録手段に記録された関係データにしたがって、質的評価値と量的評価値とにに基づくグラフを描画するグラフ化機能を実現させるためのプログラム、を附加してもよい。これにより、分析結果の把握が一層容易になる。

【0044】

上記第1又は第2の発明の検索キーワード分析プログラムにおいて、検索キーワードと質的評価値と量的評価値とが求められる場合、量的評価値は、検索キーワードの入力数が多いほど値が大きくなる検索結果のニーズ値であり、量的評価規則は、検索キーワードの入力数が多いほどニーズ値が増加する規則であるとしてもよい。

【0045】

上記第1又は第2の発明の検索キーワード分析プログラムにおいて、Webアクセスログは、WWWにおけるアクセス毎に、アクセス要求側識別データとアクセス先識別データとアクセス時刻を示すデータとを関係付けており、質的評価規則は、アクセスシーケンスにおける検索キーワード入力後のアクセスから最後のアクセスまでの時間間隔が長いほどアクセス要求側の満足度が増加する規則であり、質的評価値は、満足度であるとしてもよい。

【0046】

上記第1又は第2の発明の検索キーワード分析プログラムにおいて、質的評価規則は、分析候補の検索キーワードと分析候補の検索キーワードの入力後にア

セスされた場合に成功度を増加するアクセス先識別データとを関連付けたデータと記録手段に記録されたキーワードデータとに基づいて、アクセスシーケンスにおいて分析候補の検索キーワードが入力されておりその入力後分析候補の検索キーワードと関連付けされているアクセス先識別データの示すアクセス先がアクセスされている場合に、アクセスシーケンスに対する成功度を増加する規則であり、質的評価値は、成功度であるとしてもよい。

【0047】

この場合において、質的評価規則は、分析候補の検索キーワードが入力されてから分析候補の検索キーワードと関連付けされているアクセス先識別データの示すアクセス先がアクセスされるまでのアクセス数が少ないほど加えられる成功度を大きくするとしてもよい。

【0048】

上記第1又は第2の発明の検索キーワード分析プログラムにおいて、Webアクセスログは、アクセス先識別データの示すアクセス先がハイパリンクを使ってアクセスされた場合に、アクセス先識別データの示すアクセス先の前段でアクセスされたアクセス先を示すレファラ情報とアクセス先識別データとを関連付けたデータを含み、アクセスシーケンス抽出機能は、レファラ情報が検索システムを示す場合に、レファラ情報に関連付けされているアクセス先識別データの示すアクセス先へのアクセスを含むアクセスシーケンスを抽出し、キーワード抽出機能は、レファラ情報が検索システムを示す場合に、レファラ情報に含まれている検索キーワードを抽出し、評価値算出機能は、アクセスシーケンスに対する検索キーワード入力後の質的評価値を算出するとしてもよい。

【0049】

これにより、アクセス先で記録されたWebアクセスログのみからであっても上記と同様の効果を得ることができる。

【0050】

上記第1又は第2の発明の検索キーワード分析プログラムにおいて、検索キーワードと質的評価値と量的評価値の対応関係を示す関係データが求められる場合、Webアクセスログは、アクセス要求側識別データとアクセス先識別データと

アクセスされた時間を示す時間データとを関係付けたデータであり、評価値算出機能は、時間データに基づいて複数の期間毎の関係データを求め、コンピュータに、複数の期間毎の関係データに基づいて、検索キーワードについて質的評価値と量的評価値との遷移状況を表すトレンドデータを求め、記録手段に記録するトレンド分析機能を実現させるとしてもよい。

【0051】

これにより、WWWにおける分析対象の部分をアクセスするために検索システムに入力される検索キーワードの質的評価値及び量的評価値の変化が時系列に求められる。

【0052】

上記第1又は第2の発明の検索キーワード分析プログラムにおいて、検索キーワードと質的評価値と量的評価値の対応関係を示す関係データが求められる場合、Webアクセスログは、アクセス要求側識別データとアクセス先識別データとアクセスされた時間を示す時間データとを関係付けたデータであり、評価値算出機能は、時間データに基づいて複数の期間毎の前記関係データを求め、コンピュータに、検索キーワードと検索キーワードの属するカテゴリとを関係付けたカテゴリデータと複数の期間毎の関係データとに基づいて複数の期間毎にカテゴリについて質的評価値と量的評価値とを算出し、記録手段に記録するカテゴリ形成機能と、カテゴリについて質的評価値と量的評価値との遷移状況を表すトレンドデータを求め、記録手段に記録するトレンド分析機能とを実現させるとしてもよい。

【0053】

これにより、検索キーワードを統合して分析者の求めるカテゴリについて分析を行うことができる。

【0054】

上記の検索キーワード分析プログラム又はこのプログラムを記録した記録媒体を用いることによって、上記各機能を有していないコンピュータ、コンピュータシステム、サーバ、クライアントに対しても、簡単にこれらの機能を付加することができる。

【0055】

本発明で実現される機能と同様の動作を実現する手段を検索キーワード分析システムに具備することにより、同様の作用効果を得ることができる。

【0056】

本発明の検索キーワード分析プログラム又はこのプログラムを記録した記録媒体をコンピュータに適用することによって、入力された検索キーワードに基づく検索システムの検索結果に対する質的評価値を容易に把握するための検索キーワード分析方法がコンピュータ上で実施される。

【0057】**【発明の実施の形態】**

以下、図面を参照しながら本発明の実施の形態について説明する。なお、Webサイトとは少なくとも1つのWebページを含むWebページのグループを指す。

【0058】

以下の各図において同一の部分については同一の符号を付してその説明を省略し、異なる部分についてのみ詳しく説明する。

【0059】**(第1の実施の形態)**

本実施の形態においては、検索システムによる検索結果を表示する検索結果ページからの閲覧者の閲覧行動（2つ以上のページ遷移を含む）を分析することにより、閲覧者から入力された検索キーワードに基づく検索結果に対する質的評価値を高精度に求める検索キーワード分析システムについて説明する。

【0060】

上述したように、検索システムは、閲覧者が入力した検索キーワードに応じて、この検索キーワードと関連があると判断されたキーワード関連ページについての情報を抽出し、リスト化し、閲覧者に表示する。

【0061】

閲覧者は、検索システムに対して自己の意思に基づいて自由に検索キーワードを入力する。

【0062】

このため、各種の検索キーワードの入力頻度は、各種の検索キーワード間で一定の割合となることはなく、ある程度隔たりが発生する。

【0063】

したがって、顧客サイトをアクセスするために入力された検索キーワードによる検索結果を量的に評価する量的評価値は、検索キーワード毎に差が生じる。

【0064】

量的評価値の例としては、各検索キーワードについて検索結果のニーズの量、数、多さを表すニーズ値があり、例えば検索キーワードの入力回数をニーズ値とすることができます。

【0065】

その他にも、量的評価値の例としては、上述した検索キーワードについて抽出されたキーワード関連ページの件数、検索結果ページのクリック回数などがある。

【0066】

その一方で、検索キーワードに基づいて検索されたキーワード関連ページが必ずしも閲覧者の求めている内容とは限らない。

【0067】

したがって、顧客サイトをアクセスするために入力された検索キーワードによる検索結果を質的に評価する質的評価値を求める意義は大きい。

【0068】

質的評価値の例としては、検索キーワードによる検索結果がどの程度効果的に機能したかを表す度合いである効果度がある。さらに、この効果度の具体例としては、検索システムを利用した閲覧者が、その検索キーワードに対して検索システムが提示した情報サービスにどの程度満足したかを表す満足度がある。また、効果度の他の具体例として、Webサイト運営者がアクセスして欲しいと考えているWebページにどの程度自サイトの閲覧者がアクセスしているかを表す成功度がある。

【0069】

質的評価値として、例えば、閲覧者が検索キーワードによる検索結果をどの程度重視しているかを表す重視度もある。

【0070】

なお、質的評価値についても、検索システムの性能、リストとして表示されたキーワード関連ページの質、内容などに応じて、検索キーワード毎に差が生じる。

【0071】

質的評価値を調査することは、Webページ又はWebサイトの運営者にとって、自己の運営しているWebページ又はWebサイト、検索システムが閲覧者を例えどどの程度満足させているか把握するために有効であり、Webページ又はWebサイトのデザインやコンテンツ、検索システムの改善に有効である。

【0072】

なお、本実施の形態においては、質的評価値のうち満足度を求める場合について説明するが、満足度ではない他の質的評価値を求める場合も同様である。

【0073】

図1は、本実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図である。

【0074】

図2は、顧客サイトの構成の一例を示すブロック図である。

【0075】

記録装置4は、調査対象Webサイトの運営者である顧客5によって運営される顧客サイト6のWebアクセスログ7を記録する。

【0076】

顧客サイト6は、サイト内検索システム6aと、一以上のWebページP₁～P_nを持つ。

【0077】

サイト内検索システム6aは、ある特定のサイト内のWebページについての検索を行う。この図1の例では、サイト内検索システム6aは、顧客サイト6のWebページP₁～P_nについての検索を行う。

【0078】

一般的な閲覧者8のうち、顧客サイト6をアクセスした閲覧者を顧客サイト閲覧者8aとする。

【0079】

表1に、Webアクセスログ7に含まれる情報の一例を示す。

【0080】

【表1】

表1 Webアクセスログに含まれる情報（顧客サイト閲覧者）

ホスト名	User ID	日付	時刻	リクエスト	コード	終了 バイト数	転送 情報	ユーザエージェント
pc001.○○○. co.jp	—	2002/2/14	18:56:14	/index.html	200	400	—	MOOiiIa/4.0 (compatible:MOIE...)
ws20.△△△. co.jp	—	2002/2/14	18:57:30	/cgi/login.cgi	200	700	—	MOOiiIa/4.7 [△△△5LLL] (X11...)
pc001.○○○. co.jp	—	2002/2/14	18:58:01	/products/ index.html	200	2400	—	w△m/0.3tm□e-p21- 26tm○e-1.5.2
pc001.○○○. co.jp	—	2002/2/14	18:58:40	/products/info/ p001.pdf	200	3000	—	MOOiiIa/4.7 [△△△5LLL] (X11...)
ws20.△△△. co.jp	—	2002/2/14	18:59:10	/cgi/order.cgi	200	700	—	MOOiiIa/4.0 (compatible:MOIE...)

【0081】

Webアクセスログ7には、顧客サイト閲覧者8aによる顧客サイト6のア

セス毎に、ホスト名又はIPアドレス、顧客サイト閲覧者8aが顧客サイト6にアクセスした日付と時刻を示す情報、顧客サイト閲覧者8aの発信したリクエスト（URL）、レファラ情報、アクセスした顧客サイト閲覧者8aのID（UserID）、などから構成されるレコードが含まれる。

【0082】

なお、レファラ情報とは、ハイパーリンクを使ってあるWebページがアクセスされた場合に、このアクセスされたWebページがどのWebページからリンクを辿ってアクセスされたかを示す情報である。表1においては、顧客サイト閲覧者8aのIDとレファラ情報については、具体的な値の記述を省略している。

【0083】

本実施の形態に係る検索キーワード分析システム9は、記録媒体10に記録されている検索キーワード分析プログラム11を読み出して実行する。

【0084】

検索キーワード分析プログラム11は、システム9に読み込まれ、実行されると、ログ前処理機能12、アクセスシーケンス抽出機能13、キーワード抽出機能14、評価値算出機能37を実現する。

【0085】

また、評価値算出機能37は、満足度算出機能15、満足度・キーワード関連付け機能16を含む。

【0086】

上記各機能12～16によって作成されたデータは、記録装置38に記録される。各機能12～16は、記録装置38に記録されたデータを読み込み可能である。記録装置38としては、例えば内部メモリ等が適用される。

【0087】

ログ前処理機能12は、記録装置4に記録されたWebアクセスログ7から、検索キーワード分析システム9に必要な情報を入力する。

【0088】

そして、ログ前処理機能12は、入力した情報に基づいて、各ページアクセスに対して重複のないアクセス識別子（ViewID）を付与する。

【0089】

表2に、ログ前処理機能12によって生成された前処理データの一例を示す。

【0090】

【表2】

表2 前処理データ

ViewID	ホスト名	User ID	日付	時刻	リクエスト	ユーザーラ 情報	ユーザージェント
View1	pc001.○○○. co.jp	—	2002/2/14	18:56:14	/index.html	—	MOOiiIa/4.0 (compatible:MOIE...)
View2	ws20.△△△. co.jp	—	2002/2/14	18:57:30	/cgi/login.cgi	—	MOOiiIa/4.7 [△△5LLL] (X11...)
View3	pc001.○○○. co.jp	—	2002/2/14	18:58:01	/products/ index.html	—	w△m/0.3+m□e-p21- 26+m○e-1.5.2
View4	pc001.○○○. co.jp	—	2002/2/14	18:58:40	/products/ info/p001.pdf	—	MOOiiIa/4.7 [△△5LLL] (X11...)
View5	ws20.△△△. co.jp	—	2002/2/14	18:59:10	/cgi/order.cgi	—	MOOiiIa/4.0 (compatible:MOIE...)

【0091】

アクセスシーケンス抽出機能13は、前処理データに基づいて、顧客サイト閲

観者8aによってアクセスされた一連のWebページの列（以下、「アクセスシーケンス」という）を抽出し、各アクセスシーケンスに重複のないアクセスシーケンス識別子を付与し、アクセスシーケンス抽出データを生成する。

【0092】

アクセスシーケンスとは、同一又は同一として扱われる閲覧者、すなわち同一のアクセス要求側による連続的なアクセスをいう。

【0093】

アクセスシーケンス抽出機能13は、ホスト名又はIPアドレスとユーザエージェントとの組み合わせが一致する場合に同一のアクセス要求側として扱い、ホスト名又はIPアドレスとユーザエージェントの各組み合わせに対してアクセス要求側識別子を付与し、同一のアクセス要求側から連続的にアクセスされた一以上のWebページをアクセスシーケンスとして切り出す。

【0094】

アクセスシーケンス抽出機能13は、所定のアクセスシーケンス抽出規則にしたがって、一連のページアクセスからアクセスシーケンスを切り出し、アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを作成する。

【0095】

例えば、アクセスシーケンス抽出規則として、一般的なWebアクセスログ分析で用いられているセッションを抽出する規則を適用可能である。具体的には、ある閲覧者によるWebページのアクセスから次のWebページへのアクセスの間に30分以上の時間が経過した場合（次のWebページへの遷移がなかった場合を含む）、その間をセッションの境界（もしくはセッションの終端）とし、境界と境界との間をセッションとして抽出する。そして、このセッションをアクセスシーケンスとして用いる。

【0096】

なお、アクセスシーケンスとアクセスシーケンスとの間の境界を判定するための経過時間は、上記の30分に限定されるものではなく、自由に設定可能である。すなわち、アクセスシーケンス抽出規則は、複数のアクセスが同一のアクセス要求側から所定の時間間隔内でアクセスされている場合に複数のアクセスが同一

のアクセスシーケンスに含まれると判断する規則としてもよい。

【0097】

また、アクセスシーケンスを抽出する場合において、W e b アクセスログ7に記録された閲覧者の I D (UserID) 又はCookie情報に基づいて、アクセス要求側の同一性を判断してもよい。

【0098】

表3に、アクセスシーケンスデータの一例を示す。

【0099】

【表3】

表3 アクセスシーケンスデータ

ViewID	VisitID	VisiterID	日付	時刻	リクエスト
View1	Visit1	Visiter1	2002/2/14	18:56:14	/AAA.html
View2	Visit2	Visiter2	2002/2/14	18:57:30	/BBB.cgi
View3	Visit1	Visiter1	2002/2/14	18:58:01	/CCC.html
View4	Visit2	Visiter2	2002/2/14	18:58:40	/GGG/DDD.cgi
View5	Visit3	Visiter1	2002/2/14	19:42:10	/EEE.html

【0100】

表3のアクセスシーケンスデータには、各アクセスシーケンスを識別するためのアクセスシーケンス識別子 (VisitID) 、ホスト名とユーザエージェントの組み合わせを識別するアクセス要求側識別子 (VisitorID) などから構成されるコードが含まれる。

【0101】

キーワード抽出機能14は、ログ前処理機能12によって生成された前処理データの中から、顧客サイト閲覧者8aがサイト内検索システム6aにおいて入力した検索キーワードを抽出し、検索キーワードの抽出結果を示すキーワードデータを作成する。

【0102】

表4に、W e b アクセスログ7における検索システムへのリクエストの例を示す。

【0103】

【表4】

表4 検索システムへのリクエスト

例1	http://www.○○○.co.jp/search.cgi?q=xxxx
例2	http://www.○○○.co.jp/search.cgi?q=yyyy&page=10
例3	http://www.○○○.co.jp/search.cgi?q=%E3%82%8A%E3%82%93%E3%81%94

【0104】

表4における例1は、検索キーワード「xxxx」を含んだ一番基本的な検索システムへのリクエストの例である。

【0105】

例2では、検索キーワード「yyyy」と検索結果の10ページ目とがリクエストされている。

【0106】

例3では、検索キーワード「りんご」を示すコードがリクエストに含まれている。

【0107】

Webアクセスログ7の中から検索キーワードを抽出するためには、検索システムの場所及び検索キーワードを示すパラメータを表すクエリーパターンに関する情報を用いる。

【0108】

表5に、検索システムの場所とクエリーパターンに関する情報の一例を示す。

【0109】

【表5】

表5 検索システムの場所とクエリーパターン

検索システムの場所	クエリーパターン
http://www.○○○.co.jp/search.cgi	q

【0110】

すなわち、キーワード抽出機能14は、前処理データから検索システムの場所に関する情報に基づいて、検索システムを利用したアクセスを抽出する。

【0111】

次に、キーワード抽出機能14は、クエリーパターンに関する情報に基づいて、検索キーワードを抽出する。

【0112】

表6に、上記表4、5の例に基づいて、キーワード抽出機能14によって抽出されたパラメータの例を示す。

【0113】

【表6】

表6 抽出されたパラメータ

パラメータ	
例1	xxxx
例2	yyyy
例3	%E3%82%8A%E3%82%93%E3%81%94

【0114】

抽出されたパラメータの中には、この表6の例3に示すようにエンコードされたパラメータも存在する。

【0115】

キーワード抽出機能14は、エンコードされたパラメータについては判読可能な文字列へのデコードを行い、キーワードデータを生成する。

【0116】

表7に、キーワードデータの例を示す。

【0117】

【表7】

表7 キーワードデータ

ViewID	検索キーワード
View6	りんご
View21	apple
View43	banana
View56	砂糖
View81	カレイ

【0118】

表7のキーワードデータには、各アクセス識別子とそのアクセス識別子に対応するアクセスにおいて入力された検索キーワードとから構成されるレコードが含まれている。

【0119】

満足度算出機能15は、アクセスシーケンス抽出機能13によって生成されたアクセスシーケンスデータに基づいて、閲覧者の満足度を算出する。

【0120】

満足度を算出するために、満足度をどのように算出するかを表す質的評価規則が定義される。

【0121】

例えば、質的評価規則として、サイト内検索システム6aに検索キーワードを入力した後の顧客サイト閲覧者8aによる顧客サイト6の閲覧時間（以下、「検索後閲覧時間」という）が長いほど満足度が高くなる規則を定義する。

【0122】

表8に、検索後閲覧時間と満足度とを関係付けた例を示す。

【0123】

【表8】

表8 検索後閲覧時間と満足度

検索後閲覧時間	満足度
閲覧なし	0
1ページのみの閲覧	0
1分以下	10
1-2分	20
2-3分	30
3-5分	40
5-7分	50
7-10分	60
10-15分	70
15-20分	80
20-25分	90
30分以上	100

【0124】

検索後閲覧時間は、満足度を求めるアクセスシーケンスについて最後のアクセス時刻から最初のキーワード関連ページへのアクセス時刻を引いて得られる時間間隔とする。ただし、検索結果ページからキーワード関連ページに遷移することなくアクセスシーケンスが終わっていた場合には検索後閲覧時間を「閲覧なし」とする。また、検索結果ページからキーワード関連ページに遷移して、その後他のサイト内のWebページを閲覧することなくアクセスシーケンスが終わっている場合には、検索後閲覧時間を「1ページのみの閲覧」とする。

【0125】

すなわち、満足度算出機能15は、アクセスシーケンスデータを読み込み、アクセスシーケンスごとに検索後閲覧時間を計算する。

【0126】

次に、満足度算出機能15は、上記表8のような検索後閲覧時間と満足度の対応データを読み込み、アクセスシーケンスごとの満足度を算出する。

【0127】

表9に、満足度算出機能15によって生成された満足度算出データの例を示す。

【0128】

【表9】

表9 満足度算出データ

ViewID	VisitID	満足度
View1	Visit1	0
View2	Visit2	20
View3	Visit1	0
View4	Visit2	20
View5	Visit3	10

【0129】

表9の満足度算出データには、各アクセス識別子とそのアクセスの属するアクセスシーケンス識別子に対応する満足度とから構成されるレコードが含まれている。

【0130】

満足度・キーワード関連付け機能16は、満足度算出機能15で生成された満足度算出データと、キーワード抽出機能14で生成されたキーワードデータを統合し、検索キーワードごとの満足度を算出し、満足度と検索キーワードとを関係付け、関係データを作成する。

【0131】

満足度・キーワード関連付け機能16は、まず、満足度算出データとキーワードデータを読み込む。

【0132】

次に、満足度・キーワード関連付け機能16は、アクセス識別子に基づいて、満足度算出データとキーワードデータを統合する。

【0133】

表10に、満足度算出データとキーワードデータの統合結果を示す。

【0134】

【表10】

表10 満足度算出データとキーワードデータの統合結果

ViewID	満足度	検索キーワード
View6	10	りんご
View21	0	apple
View83	30	りんご
View121	20	砂糖
View137	0	apple

【0135】

そして、満足度・キーワード関連付け機能16は、検索キーワードごとに満足度の平均を求め、関係データを作成する。

【0136】

表11に、関係データの例を示す。

【0137】

【表11】

表11 関係データ

検索キーワード	満足度
りんご	23.4
apple	17.8
砂糖	2.4
肉野菜	2.4
蜂蜜	11.6

【0138】

図3は、本実施の形態に係る検索キーワード分析システム9の動作の一例を示すフローチャートである。

【0139】

ステップS1において、ログ前処理機能12は、Webアクセスログ7に含まれる情報から必要な情報を入力し、前処理データを作成する。

【0140】

ステップS2において、アクセスシーケンス抽出機能13は、前処理データとアクセスシーケンス抽出規則とに基づいて、アクセスシーケンスを抽出し、アクセスシーケンスデータを作成する。

【0141】

ステップS3において、満足度算出機能15は、アクセスシーケンスデータと満足度算出規則とに基づいて、アクセスシーケンス毎の満足度を算出し、満足度算出データを作成する。

【0142】

ステップS4において、キーワード抽出機能14は、前処理データに基づいて、サイト内検索システム6aに入力された検索キーワードを抽出し、キーワードデータを作成する。

【0143】

ステップS5において、満足度・キーワード関連付け機能16は、満足度算出データと、キーワードデータに基づいて、検索キーワードごとの満足度を算出し、関係データを作成する。

【0144】

なお、ステップS4は、ステップS2及びステップS3の前に実行されてもよく、後に実行されてもよく、並列に実行されてもよい。

【0145】

図4は、本実施の形態に係るキーワード抽出機能14の動作の一例を示すフローチャートである。

【0146】

ステップT1において、キーワード抽出機能14は、検索システムの場所とクエリーパターンに関する情報を読み込む。

【0147】

ステップT2において、キーワード抽出機能14は、ログ前処理機能12によって前処理されたWebアクセスログ7から一つのレコードを読み込む。

【0148】

ステップT3において、キーワード抽出機能14は、読み込んだレコードに含まれているリクエストが検索システムの場所を含むか判定する。

【0149】

リクエストが検索システムの場所を含むと判定された場合、ステップT4において、キーワード抽出機能14は、検索システムの場所に対応するクエリーパターンに基づいて、リクエストからパラメータを抽出する。

【0150】

リクエストが検索システムの場所を含むと判定されなかった場合、キーワード抽出機能14は、ステップT2以降の処理を繰り返す。

【0151】

ステップT5において、キーワード抽出機能14は、抽出したパラメータに対してデコードが必要か否か判定する。

【0152】

デコードが必要な場合、ステップT6において、キーワード抽出機能14は、抽出したパラメータをデコードする。

【0153】

ステップT 7において、キーワード抽出機能14は、ログ前処理機能12によって前処理されたWebアクセスログ7から終端のレコードを読み込んだか否か判定する。

【0154】

終端のレコードが読み込まれていないと判定された場合、キーワード抽出機能14は、ステップT 2以降の処理を繰り返す。

【0155】

終端のレコードが読み込まれたと判定された場合、ステップT 8において、キーワード抽出機能14は、抽出されたパラメータを検索キーワードとし、キーワードデータを作成する。

【0156】

図5は、本実施の形態に係る満足度算出機能15の動作の一例を示すフローチャートである。

【0157】

ステップU 1において、満足度算出機能15は、アクセスシーケンス抽出機能13によって生成されたアクセスシーケンスデータを読み込む。

【0158】

ステップU 2において、満足度算出機能15は、アクセスシーケンス毎に、検索後閲覧時間を計算する。

【0159】

ステップU 3において、満足度算出機能15は、検索後閲覧時間と満足度とを関係付けたデータを読み込む。

【0160】

ステップU 4において、満足度算出機能15は、アクセスシーケンス毎に、検索後閲覧時間に対応する満足度を求める。

【0161】

ステップU 5において、満足度算出機能15は、アクセスシーケンス毎の満足度を示す満足度算出データを作成する。

【0162】

なお、ステップU3は、ステップU1及びステップU2の前に実行されてもよく、並列に実行されてもよい。

【0163】

以上説明した本実施の形態においては、検索キーワードについての分析結果として、顧客サイト閲覧者8aがその検索キーワードによる検索結果についてどれほど満足したか求めることができ、検索キーワードに対する検索結果について量的評価値のみではなく、十分な精度で質的評価値を分析することができる。

【0164】

本実施の形態により、例えば、顧客5は、検索結果に対して顧客サイト閲覧者8aがどの程度満足しているかを定量的に知ることができる。例えば、上記表11の結果より、検索キーワード「りんご」による検索結果については顧客サイト閲覧者8aの満足度が高いのに対し、検索キーワード「砂糖」又は検索キーワード「肉野菜」による検索結果については顧客サイト閲覧者8aの満足度が低いことが分かる。

【0165】

なお、質的評価値としては、満足度等に代表される効果度の他に、例えば重視度や利便度も考えられる。

【0166】

重視度とは、Webサイトの閲覧者が検索キーワードに関連する事柄について、どの程度重視しているかを表す指標である。

【0167】

利便度とは、Webサイトの閲覧者が検索キーワードに関連する事柄について、どの程度容易に閲覧やWebサイトへの書き込み等の操作ができたかを表す指標である。

【0168】

重視度や利便度も、満足度と同様に検索後閲覧時間によって求めることが可能である。

【0169】

検索後閲覧時間を満足度、重視度、利便度のいずれを算出するのに利用するか

は、分析対象となるWebサイトや閲覧者の性質、分析者の観点によって適宜選択できる。

【0170】

例えば、商品を販売しているECサイトにおいて、他にも類似の競合サイトが多く存在するケースでは、顧客は検索キーワードに関連したコンテンツが顧客の期待にそわなかつた場合に、比較的短時間で閲覧をあきらめてサイトを去り、別のサイトを閲覧することが考えられる。このようなケースでは、検索後閲覧時間は満足度としての利用が適切である。例えば、「蜂蜜」という検索キーワードに対する検索後閲覧時間が短いために、「蜂蜜に関する満足度が低い」という分析結果が得られた場合、この結果は「蜂蜜」に関連するコンテンツを改良して、満足度を向上させるような対策を打つ必要性が高い、という判断をする根拠として有効である。

【0171】

一方、他に類似の競合サイトが少ない場合や、提供しているコンテンツが閲覧者の嗜好を満たすような内容というよりはむしろ合理的必要性に応じて閲覧している場合や、その検索キーワードに関する情報をできるだけ当該Webサイトから得たいと思っている閲覧者が多数であるような場合においては、閲覧者が検索キーワードに関連する内容を重視している度合いに応じて、長時間サイトに滞在して、その情報を閲覧すると考えられる。したがって、検索後閲覧時間を重視度の算出に利用することが適切な場合がある。この場合に、ある検索キーワードに対する検索後閲覧時間が短いために、それに関する重視度が低いという分析結果が得られた場合、この結果はそのコンテンツについては以後それほど改良のための対策を打つ必要性が低い、という判断をする根拠として有効である。

【0172】

さらに、利便度においては、例えば提供しているコンテンツをどれだけ閲覧者が容易に理解したか、情報の入力がどの程度容易であったかを評価する指標として有効である。

【0173】

例えば検索キーワードに応じてアンケート内容を変えているWebサイトにお

いて、検索後閲覧時間が長いということは、アンケートの返答の入力に手間取っていると判断され、利便度が低いと判断される。

【0174】

このように利便度の低いコンテンツがあれば、利便度を向上させるべく、対策を打つ必要性が高い、という判断をする根拠として有効である。

【0175】

(第2の実施の形態)

本実施の形態においては、上記第1の実施の形態に係る検索キーワード分析システムの変形例について説明する。なお、本実施の形態においては、質的評価値のうち満足度を求め、量的評価値のうちニーズ値を求める場合について説明するが、他の質的評価値及び量的評価値を求める場合も同様である。

【0176】

図6は、本実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図である。

【0177】

本実施の形態に係る検索キーワード分析システム18は、記録媒体19に記録されている検索キーワード分析プログラム20を読み出して実行する。

【0178】

検索キーワード分析プログラム20は、システム18に読み込まれ、実行されると、ログ前処理機能12、アクセスシーケンス抽出機能13、キーワード抽出機能14、評価値算出機能39を実現する。

【0179】

また、評価値算出機能39は、満足度算出機能15、ニーズ値算出機能21、満足度・ニーズ値関連付け機能22を含む。

【0180】

ニーズ値算出機能21、満足度・ニーズ値関連付け機能22によって作成されたデータは、記録装置38に記録される。ニーズ値算出機能21、満足度・ニーズ値関連付け機能22は、記録装置38に記録されたデータを読み込み可能である。

【0181】

ニーズ値算出機能21は、キーワード抽出機能14によって生成されたキーワードデータに基づいて、検索キーワードごとのニーズ値を算出する。

【0182】

ニーズ値を算出するために、ニーズ値をどのように算出するかを表す量的評価規則が定義される。

【0183】

例えば、量的評価規則として、検索キーワードの入力総数に対する各検索キーワードの占める割合を百分率で表した値をニーズ値と定義する。

【0184】

ニーズ値算出機能21は、キーワード抽出機能14によって生成されたキーワードデータを読み込み、検索キーワードごとの入力個数をカウントする。

【0185】

次に、ニーズ値算出機能21は、検索キーワードごとの入力個数から全検索キーワードの入力個数の合計を算出する。

【0186】

次に、ニーズ値算出機能21は、全検索キーワードの入力個数と検索キーワードごとの入力個数に基づいて、全検索キーワードの入力に対する各検索キーワードの占める割合を百分率で表す。

【0187】

そして、ニーズ値算出機能21は、検索キーワードごとのニーズ値を表すニーズ値算出データを作成する。

【0188】

表12に、ニーズ値算出データの一例を示す。

【0189】

【表12】

表12 ニーズ値算出データ

検索キーワード	ニーズ値
りんご	21.5
apple	14.6
砂糖	3.6
肉野菜	17.6
蜂蜜	2.7

【0190】

満足度・ニーズ値関連付け機能22は、満足度算出機能15によって生成された満足度算出データとニーズ値算出機能21によって生成されたニーズ値算出データを読み込む。

【0191】

そして、満足度・ニーズ値関連付け機能22は、満足度算出データとニーズ値算出データを検索キーワード毎に統合し、検索キーワードごとのニーズ値、満足度を表す関係データを作成する。

【0192】

表13に、関係データの一例を示す。

【0193】

【表13】

表13 関係データ

検索キーワード	ニーズ値	満足度
りんご	21.5	23.4
apple	14.6	17.8
砂糖	3.6	2.4
肉野菜	17.6	2.4
蜂蜜	2.7	11.6

【0194】

以上説明した本実施の形態において、顧客5は、質的評価値（例えば満足度）と量的評価値（例えばニーズ値）とを参照、比較しつつ、早急に対策を取る必要がある検索キーワードについて知ることができる。

【0195】

例えば、顧客5は、上記表13の例を参照した場合に、検索キーワード「砂糖」と「肉野菜」は同程度の満足度であり他の検索キーワードに比べて満足度が低いが、「肉野菜」は「砂糖」に比べてニーズ値が高いと解釈できる。

【0196】

この結果、顧客5は、検索キーワード「砂糖」について満足度が低いがあまり検索のニーズがなく（本実施形態の場合、検索キーワード「砂糖」による検索結果ページの閲覧者が少なく）、比較的影響力が低いと判断できる。

【0197】

これに対し、顧客5は、検索キーワード「肉野菜」は「砂糖」と同じ程度で満足度が低いが検索のニーズがあり（本実施形態の場合、検索キーワード「肉野菜」による検索結果ページの閲覧者が多く）、影響力が大きいと判断できる。

【0198】

このため、本実施の形態において、顧客5は、「肉野菜」関連のコンテンツを増やすなどの対策を優先的に講じる必要があると判断でき、迅速かつ正確に閲覧者の希望に合うようにサイト内検索システム6a、WebページP₁～P_nを変更することができる。

【0199】

なお、本実施の形態において、ニーズ値の算出方法は、キーワード抽出機能14、ニーズ値算出機能21の設定を変更することにより、自由に変更できる。例えば、検索キーワードが入力されて得られたキーワード関連ページをアクセスした顧客サイト閲覧者8aの人数又はキーワード関連ページのアクセス数に基づいてニーズ値を算出してもよい。また、例えば、アクセスシーケンスにしたがったアクセスを行った閲覧者の人数、アクセスシーケンスにしたがったアクセスの実行された回数（閲覧回数）に基づいてニーズ値を算出してもよい。

【0200】

なお、量的評価値の例として、ニーズ値の他に機会値もある。機会値とは、Webサイトがその目的を遂行するに当たって、検索キーワードに関連する観点で、どれだけの事業機会を得たかを示す指標である。

【0201】

例えば、食品を販売しているECサイトにおいて、「砂糖」という検索キーワードによって閲覧してきた閲覧者が多数いれば、それだけ多くの砂糖を販売する事業機会を得たことになる。

【0202】

また、検索キーワードごとの入力回数のような、ニーズ値の算出に用いられた数量に基づいて機会値を算出してもよい。

【0203】

検索キーワードごとの入力回数をニーズ値とするか機会値とするかについては、Webサイトや閲覧者の特性、分析者の観点に依存する。

【0204】

例えば、商品そのものを販売するというよりは、商品に関する情報を提供し、その反応を知るために運営されているWebサイトにおいては、検索キーワードごとの入力回数をニーズ値として取り扱うことが適切である。

【0205】

一方、商品そのものをオンラインで販売することを主たる目的としているWebサイトでは検索キーワードごとの入力回数を機会値として取り扱い、各検索キーワードに関連した商品の販売の機会の量を示す指標とすることが適切である。

【0206】

(第3の実施の形態)

本実施の形態においては、上記第2の実施の形態に係る検索キーワード分析システムの変形例について説明する。

【0207】

図7は、本実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図である。

【0208】

本実施の形態に係る検索キーワード分析システム24は、記録媒体25に記録されている検索キーワード分析プログラム26を読み出して実行する。

【0209】

検索キーワード分析プログラム26は、システム24に読み込まれ、実行されると、ログ前処理機能12、アクセスシーケンス抽出機能13、キーワード抽出機能14、評価値算出機能39に加えて、グラフ化機能27を実現する。

【0210】

グラフ化機能27は、満足度・ニーズ値関連付け機能22によって生成された関係データを読み込む。

【0211】

そして、グラフ化機能27は、検索キーワードごとのニーズ値・満足度の値をグラフで表現した出力データ28を作成する。

【0212】

グラフ化機能27は、所定のグラフ化規則とグラフ化に利用されるパラメータとにしたがって、検索キーワードごとに満足度とニーズ値とが関連付けされた関係データをグラフ化する。

【0213】

例えば、グラフ化機能27は、散布図を作成する。この場合、満足度とニーズ値とをパラメータとし、散布図の一方の軸（X軸）をニーズ値、他方の軸（Y軸）を満足度とし、関係データにおけるニーズ値・満足度の値に基づいて各検索キーワードを表す点をプロットする。

【0214】

図8は、検索キーワード毎の満足度とニーズ値とを表すグラフの一例である。

【0215】

本実施の形態においては、検索キーワードごとの満足度とニーズ値とがグラフ化されて表示されるため、顧客5に視覚的な分析結果が提供される。

【0216】

したがって、顧客5は、より人間が直感的に解釈・判断しやすい分析結果を得ることができる。

【0217】

例えば、顧客5は、上記図8のグラフから、検索キーワード「りんご」についてはニーズ値が高く満足度も高いのに対して、検索キーワード「スーパー」「肉

野菜」についてはニーズ値が高いにもかかわらず満足度が低く、早急にサイト内検索システム6a、WebページP₁～P_nを変更し、対策を取る必要があると解釈される。

【0218】

なお、本実施の形態におけるグラフ化機能27の設定を変更することにより、グラフの軸の値を変換でき、グラフ表示効果を向上させることができる。例えば、グラフ化機能27は、ニーズ値、満足度を0から1の間に正規化するとしてもよい。また、ニーズ値、満足度に対数処理を施してグラフ表示してもよい。さらに、グラフ化機能27は、ニーズ値、満足度を棒グラフで表してもよい。さらに、グラフ化機能27は、グラフに適切な補助線を加えてもよい。

【0219】

(第4の実施の形態)

本実施の形態においては、上記第1乃至第3の実施の形態に係る検索キーワード分析システムの変形例について説明する。なお、以下においては、上記第1の実施の形態についての変形例を説明するが、他の第2、第3の実施の形態についても同様の変形が可能である。

【0220】

なお、本実施の形態においては、質的評価値のうち成功度を求める場合について説明するが、成功度ではない他の質的評価値を求める場合も同様である。

【0221】

図9は、本実施の形態に係る検索キーワード分析システムの一例を示すブロック図である。

【0222】

本実施の形態に係る検索キーワード分析システム29は、記録媒体30に記録されている検索キーワード分析プログラム31を読み出して実行する。

【0223】

検索キーワード分析プログラム31は、システム29に読み込まれ、実行されると、ログ前処理機能12、アクセスシーケンス抽出機能13、キーワード抽出機能14、評価値算出機能40を実現する。

【0224】

また、評価値算出機能40は、成功度算出機能32、成功度・キーワード関連付け機能41を含む。

【0225】

成功度算出機能32は、質的評価規則にしたがって、アクセスシーケンス抽出機能13によって算出されたアクセスシーケンスデータとキーワード抽出機能14によって算出されたキーワードデータとに基づいて、成功度を算出する。

【0226】

本実施の形態に係る質的評価規則として、例えば以下のようない内容の規則が定義される。

【0227】

顧客5は、代表的な検索キーワード（分析候補の検索キーワード）に対して、閲覧者にアクセスして欲しいWebページ（以下、「アクセス希望ページ」という）の識別情報を関連付けたリスト（以下、「アクセス希望ページリスト」という）を作成する。

【0228】

そして、顧客サイト閲覧者8aが検索結果ページ上のリンクにしたがってアクセス希望ページをアクセスしたら成功度をカウントアップする。

【0229】

ただし、検索結果ページからアクセス希望ページまでのリンククリック数が増加するにつれて成功度を減少させる。

【0230】

なお、顧客5が指定しなかった検索キーワードは「（その他）」とし、成功度は一切加算せず評価対象外とする。

【0231】

以上述べた質的評価規則において利用されるアクセス希望ページリストの一例を、表14に示す。

【0232】

【表14】

表14 アクセス希望ページリスト

検索キーワード	アクセス希望ページ
りんご	/apple/index.html
apple	/apple/index.html
肉野菜	/meat/index.html
肉野菜	/vegetable/index.html
...	...
(その他)	

【0233】

なお、一つの検索キーワードに対して複数のアクセス希望ページが対応付けされていてもよく、複数の検索キーワードが一つのアクセス希望ページに対応付けされていてもよい。

【0234】

また、表15に、検索結果ページからアクセス希望ページまでのリンククリック数が増加するにつれて成功度を減少させるための成功度重み付け表の一例を示す。

【0235】

【表15】

表15 成功度重み付け規則

検索結果ページ からのクリック数	成功度
1	100
2	80
3	60
4	40
5以上	20
到達せず	0

【0236】

成功度算出機能32は、アクセスシーケンス抽出機能13によって生成されたアクセスシーケンスデータと、キーワード抽出機能14によって生成されたキーワードデータと、アクセス希望ページリストと、成功度重み付け表とを読み込む

。

【0237】

なお、成功度算出機能32は、キーワードデータの全てを読み込むのではなく、キーワードデータのうち検索キーワードとアクセス識別子の組み合わせのみを読み込むとしてもよい。

【0238】

次に、成功度算出機能32は、キーワードデータから、アクセス希望ページリストの検索キーワードに含まれている検索キーワードとそのアクセス識別子(ViewID)を取り出す。

【0239】

次に、成功度算出機能32は、アクセスシーケンスデータにおいて、上記作業によって取り出されたアクセス識別子と関連付けされているアクセスシーケンス識別子を取り出す。

【0240】

次に、成功度算出機能32は、取り出されたアクセスシーケンス識別子に関連するアクセス識別子とリクエストとをアクセスシーケンスデータから取り出し、リクエストに含まれている検索キーワードを取り出す。

【0241】

表16に、この段階における第1の中間生成データを示す。

【0242】

【表16】

表16 中間生成データ（1）

ViewID	VisitID	リクエスト	検索キーワード
View20	Visit6	/index.html	
View21	Visit6	/search.cgi	
View22	Visit6	/search.cgi?q=%E3%82%8A%E3%82%93%E3%81%94	りんご
View26	Visit6	/general/index.html	
View28	Visit6	/apple/index.html	
View37	Visit6	/apple/tsugaru.html	
View45	Visit7	/index.html	
View47	Visit7	/search.cgi	
View48	Visit7	/search.cgi?q=apple	apple
...

【0243】

次に、成功度算出機能32は、各検索キーワードに対するアクセス希望ページがその検索キーワードが入力されたアクセスシーケンスにおいてアクセスされているかを調査する。

【0244】

アクセス希望ページがアクセスされている場合、成功度算出機能32は、検索キーワードの入力から何クリック後にアクセス希望ページがアクセスされたかを調査する。

【0245】

なお、検索キーワードの入力前にアクセス希望ページがアクセスされていても、そのアクセスは成功度の算出において考慮しない。

【0246】

上記表16の例において、アクセス識別子「View22」に関連する検索キーワード「りんご」に対するアクセス希望ページ識別子は、「/apple/index.html」である。このアクセス希望ページ識別子「/apple/index.html」は、アクセス識別子「View22」と同一のアクセスシーケンス識別子「Visit6」と関連するアクセス識別子「View28」に関連付けされている。

【0247】

そして、アクセス希望ページ識別子「/apple/index.html」は、検索キーワード「りんご」が入力された2クリック後に、入力されている。

【0248】

次に、成功度算出機能32は、成功度と検索結果ページからのクリック数との関係に基づいて、成功度を算出する。

【0249】

上記の例及び上記表15を適用した場合、アクセスシーケンス識別子「Visit6」で表されるアクセスシーケンスでは、検索キーワード「りんご」に対して満足度「80」が与えられる。

【0250】

表17に、この段階における第2の中間生成データを示す。

【0251】

【表17】

表17 中間生成データ (2)

VisitID	検索キーワード	成功度
Visit6	りんご	80
Visit7	apple	100
Visit21	肉野菜	0
Visit23	スーパー	60
Visit30	りんご	20

【0252】

そして、成功度算出機能32は、検索キーワードごとに、成功度の平均値を求めた成功度算出データを作成する。

【0253】

成功度・キーワード関連付け機能41は、成功度算出機能32で生成された成功度算出データと、キーワード抽出機能14で生成されたキーワードデータを統合し、検索キーワードごとの成功度を算出し、成功度と検索キーワードとを関係付け、関係データを作成する。

【0254】

関係データは、例えば、項目「満足度」が項目「成功度」に変化する点を除い

て上記表11と同様の形式となる。

【0255】

以上説明した本実施の形態における成功度算出方法を用いることで、ある検索キーワードに対し顧客にアクセスしてもらいたいとと考えているWebページへ適切に誘導がなされているかを知ることができる。

【0256】

(第5の実施の形態)

本実施の形態においては、上記第1から第4の実施の形態に係る検索キーワード分析システムの変形例について説明する。

【0257】

本実施の形態においては、検索システムが顧客サイト外にある場合の検索キーワード分析システムについて説明する。

【0258】

なお、以下においては、上記第1の実施の形態についての変形例を説明するが、他の第2から第4の実施の形態についても同様の変形が可能である。

【0259】

図10は、本実施の形態に係る顧客サイトの構成の一例を示す図である。

【0260】

顧客サイト61は一以上のWebページP₁～P_nを持つ。また、顧客サイト61におけるいくつかのWebページは、顧客5以外の第3の運営主体によって運営されている複数のサイト外の検索システム（以下、「サイト外検索システム」という）341～34nの検索対象となっている。

【0261】

サイト外検索システム341～34nは、検索キーワードを入力すると、キーワード関連ページへのリンクを含む検索結果ページを表示する。

【0262】

本実施の形態においては、顧客サイト61と検索システム341～34nの運営主体が異なるため、上記第1から第4の実施形態の場合と異なり検索システム341～34nのWebアクセスログを取得することは困難である。

【0263】

したがって、本実施の形態においては、検索システム341～34nのWebアクセスログではなく、顧客サイト61のWebアクセスログに含まれるレファラ情報をを利用して検索キーワード分析を行う。

【0264】

表18に、顧客サイト61のWebアクセスログの一例を示す。

【0265】

【表18】

表18 レファラ付Webアクセスログ

ホスト名	User ID	日付	時刻	リクエスト	終了コード	転送バイト数	レファラ情報	ユーザエージェント
pc001.○○○.co.jp	—	2002/2/14	18:56:14	/index.html	200	400	http://www.○○○.ne.jp/□□□...	MOOillia/4.0 (compatible:MOIE...)
ws20.△△△.co.jp	—	2002/2/14	18:57:30	/cgi/login.cgi	200	700	—	MOOillia/4.7 [△△5LLL] (X11...)
pc001.○○○.co.jp	—	2002/2/14	18:58:01	/products/index.html	200	2400	http://www.×××.com/△△△...	w△m/0.3tm□e-p21-26tm○e-1.5.2
pc001.○○○.co.jp	—	2002/2/14	18:58:40	/products/info/p001.pdf	200	3000	http://www.○○○.ne.jp/○○○...	MOOillia/4.7 [△△5LLL] (X11...)
ws20.△△△.co.jp	—	2002/2/14	18:59:10	/cgi/order.cgi	200	700	http://www.○○○.ne.jp/□□□...	MOOillia/4.0 (compatible:MOIE...)

【0266】

図11は、サイト外検索システム341から顧客サイト61への閲覧遷移の一例を示すブロック図である。

【0267】

閲覧者によって検索キーワード入力画面35に検索キーワードが入力されると、検索結果ページ36が作成される。この検索結果ページ36を示すURLには、検索キーワード入力画面35に入力された検索キーワードが含まれる場合がある。

【0268】

閲覧者によって検索結果ページ36のリンクがクリックされ、顧客サイト61のキーワード関連ページが閲覧された場合、検索キーワードを含む検索結果ページのURLであるレファラ情報とアクセスされた顧客サイト61のキーワード関連ページのURLであるリクエストとが関連付けされて顧客サイト61のWebアクセスログに記録される。

【0269】

本実施の形態に係る検索キーワード分析システムは、上記図1に示す構成と同様の構成を持つため、ここでは図1を参照して本実施の形態の特徴的な部分について説明する。

【0270】

本実施の形態に係る顧客サイト61のWebアクセスログには、レファラ情報が含まれる。

【0271】

本実施の形態に係る検索キーワード分析システムにおいては、アクセスシーケンス抽出機能、キーワード抽出機能、満足度抽出機能の動作が、上記第1の実施の形態に係る検索キーワード分析システム9と異なる。

【0272】

本実施の形態に係るアクセスシーケンス抽出機能は、アクセスシーケンス識別子の付与、アクセスシーケンスの区切りの判定などについては、上記第1の実施の形態に係るアクセスシーケンス抽出機能13と同様の処理を実行する。

【0273】

また、本実施の形態に係るアクセスシーケンス抽出機能は、各アクセスについてのレファラ情報に基づいて、レファラ情報がサイト外検索システムを示すか否か判定する。

【0274】

表19に、サイト外検索システムを表すURLパターンと検索キーワードを抽出するためのクエリーパターンとを関係付けたデータを示す。

【0275】

【表19】

表19 パターン

検索システムのURL	クエリーパターン
http://www.△△△.com/	keyword
http://www.□□□.ne.jp/	kw
http://www.○○○.co.jp/	query
http://www.×××.co.jp/	word
http://www.○○○.com/	q

【0276】

本実施の形態に係るアクセスシーケンス抽出機能は、サイト外検索システムからのアクセスか否かを判定する場合に、各アクセスのレファラ情報が検索システムのURLパターンを含むか調査する。

【0277】

そして、本実施の形態に係るアクセスシーケンス抽出機能は、レファラ情報がサイト外検索システムを示すアクセスに対してレファラフラグを付与したアクセスシーケンスデータを作成する。

【0278】

表20に、本実施の形態に係るアクセスシーケンスデータの一例を示す。

【0279】

【表20】

表20 アクセスシーケンスデータ

ViewID	VisitID	VisitorID	日付	時刻	リクエスト	レファラ フラグ
View1	Visit1	Visitor1	2002/2/14	18:56:14	/AAA.html	1
View2	Visit2	Visitor2	2002/2/14	18:57:30	/BBB.cgi	0
View3	Visit1	Visitor1	2002/2/14	18:58:01	/CCC.html	0
View4	Visit2	Visitor2	2002/2/14	18:58:40	/GGG/DDD.cgi	0
View5	Visit3	Visitor1	2002/2/14	19:42:10	/EEE.html	1

【0280】

表20において、レファラフラグ「0」は、サイト外検索システムからのアクセスでない旨を示す。

【0281】

一方、レファラフラグ「1」は、サイト外検索システムからのアクセスである旨を示す。

【0282】

このレファラフラグを記録することにより、アクセスシーケンスの途中で検索が行われた場合であっても、検索後のアクセス状況を調査し、満足度を算出できる。

【0283】

本実施の形態に係るキーワード抽出機能は、ログ前処理機能12によって生成された前処理データに基づいて、レファラ情報が検索システムのURLパターンを含むアクセスを求める。

【0284】

そして、本実施の形態に係るキーワード抽出機能は、URLパターンが含まれているレファラ情報の中から、その検索システムのURLパターンに対応付けられたクエリーパターンに基づいて、閲覧者がサイト外検索システムにおいて入力した検索キーワードを抽出する。

【0285】

表21に、本実施の形態に係るキーワード抽出機能によって生成されたキーワ

ードデータの一例を示す。

【0286】

【表21】

表21 キーワードデータ

ViewID	レファラ情報に含まれる キーワード
View6	りんご
View21	apple
View43	banana
View56	砂糖
View81	カレイ

【0287】

本実施の形態に係る満足度算出機能は、アクセスシーケンス抽出機能によって生成されたアクセスシーケンスデータから満足度を算出する。

【0288】

満足度の算出方法としては、例えば上記第1の実施の形態で説明した規則を適用する。

【0289】

すなわち、本実施の形態に係る満足度算出機能は、本実施の形態に係るアクセスシーケンス抽出機能によって生成されたアクセスシーケンスデータを読み込んだ後、Webページのアクセスごとに検索後閲覧時間を計算する。

【0290】

検索後閲覧時間は、同一のアクセスシーケンス識別子で表されるアクセスの中で、検索システムを利用したレファラ情報を持つアクセス（上記表20においてフラグ「1」のアクセス）から最後に閲覧されたアクセスまでの時刻の差分を求めることで計算できる。

【0291】

なお、検索システムをレファラ情報として持つアクセスがないアクセスシーケンスに関しては検索後閲覧時間を記録しない。

【0292】

また、本実施の形態に係る満足度算出機能は、上記表8のような検索後閲覧時間と満足度とを関連付けたデータを読み込み、アクセスごとの満足度を算出する。

【0293】

ただし、検索後閲覧時間の記録がない場合は満足度を記録しない。

【0294】

そして、本実施の形態に係る満足度算出機能は、アクセス識別子毎にアクセスシーケンス識別子と満足度とを関連付け、上記表9と同様な満足度算出データを作成する。

【0295】

以上説明した本実施の形態においては、顧客サイト61が検索システムを具備していない場合であっても、検索キーワードの満足度を算出することができる。

【0296】

したがって、検索キーワードの分析結果を得ることができる顧客サイトを大幅に増加することができる。

【0297】

(第6の実施の形態)

本実施の形態においては、第1の実施の形態に係る検索キーワード分析システムの変形例について説明する。

【0298】

第1から第5の実施の形態においては、アクセスシーケンスの抽出や質的評価値の算出と、検索キーワードの抽出を並列に実行し、最終的に関連付け機能によって統合する方法を用いていた。

【0299】

これに対し、本実施の形態では、検索キーワードの抽出処理と質的評価値の算出処理を直列に行う。

【0300】

なお、本実施の形態では、サイト内検索について分析を行う場合について記述しているが、本実施の形態と第5の実施形態とを組み合わせることでサイト外検

索についての分析にも適用可能である。

【0301】

図12は、本実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図である。

【0302】

本実施の形態に係る検索キーワード分析システム42は、記録媒体43に記録されている検索キーワード分析プログラム44を読み出して実行する。

【0303】

検索キーワード分析プログラム44は、システム42に読み込まれ、実行されると、ログ前処理機能12、アクセスシーケンス抽出機能13、キーワード抽出機能45、評価値算出機能46を実現する。

【0304】

キーワード抽出機能45、評価値算出機能46によって作成されたデータは、記録装置38に記録される。キーワード抽出機能45、評価値算出機能46は、記録装置38に記録されたデータを読み込み可能である。

【0305】

キーワード抽出機能45は、アクセスシーケンス抽出機能13によって生成されたデータから、顧客サイト閲覧者8aがサイト内検索システム6aにおいて入力した検索キーワードを抽出する。検索キーワードの抽出には上記第1の実施の形態と同等の方法を用いる。

【0306】

すなわち、キーワード抽出機能45は、前処理データから検索システムの場所に関する情報に基づいて、検索システムに対するアクセスを抽出する。

【0307】

次に、キーワード抽出機能45は、クエリーパターンに関する情報に基づいて、検索キーワードを抽出する。なお、エンコードされたパラメータについては判読可能な文字列へのデコードを行う。

【0308】

そして、キーワード抽出機能45は、アクセスシーケンス・キーワードデータ

を生成する。この際、検索システムを利用していないアクセスに関してはキーワード項目に何も記録しない。

【0309】

表22に、アクセスシーケンス・キーワードデータの具体例を示す。

【0310】

【表22】

表22 アクセスシーケンス・キーワードデータ

ViewID	VisitID	VisitorID	日付	時刻	検索キーワード
View1	Visit1	Visitor1	2002/2/14	18:56:14	
View2	Visit2	Visitor2	2002/2/14	18:57:30	りんご
View3	Visit1	Visitor1	2002/2/14	18:58:01	
View4	Visit2	Visitor2	2002/2/14	18:58:40	
View5	Visit3	Visitor1	2002/2/14	19:42:10	たまねぎ

【0311】

評価値算出機能46は、アクセスシーケンス・キーワードデータに基づいて、質的評価値、例えば満足度を算出する。質的評価規則は第1の実施の形態の同様の規則を用いる。

【0312】

すなわち、評価値算出機能46は、アクセスシーケンス・キーワードデータを読み込み、アクセスシーケンスごとに検索後閲覧時間を計算する。検索後閲覧時間は、同一のアクセスシーケンス識別子によるアクセスの中で、キーワード関連ページへのアクセスと一番新しいアクセスの時刻の差分から計算される。

【0313】

次に、評価値算出機能46は、上記表8のような検索後閲覧時間と満足度の対応データを読み込み、アクセスシーケンスごとの満足度を算出する。

【0314】

表23に、この段階における中間生成データを示す。

【0315】

【表23】

表23 中間生成データ

ViewID	VisitID	満足度	検索キーワード
View1	Visit1	0	
View2	Visit2	20	りんご
View3	Visit1	0	
View4	Visit2	20	
View5	Visit3	10	たまねぎ

【0316】

次に、評価値算出機能46は、アクセスシーケンスについて満足度と検索キーワードを求める。この際、検索キーワードの記録のないアクセスシーケンスは処理対象としない。

【0317】

そして、評価値算出機能46は、検索キーワードごとに、満足度の平均を求め、関係データを生成する。

【0318】

関係データの具体的な内容は、上記表11と同様である。

【0319】

なお、本実施の形態において、評価値算出機能46は、量的評価値、例えばニーズ値を算出するとしてもよい。このニーズ値の算出方法は、上記第2の実施の形態で説明した方法が適用できる。また、評価値算出機能46によって質的評価値と量的評価値とを算出した場合、上記第3の実施の形態と同様にグラフ化機能27によって関係データに基づくグラフを表示するとしてもよい。

【0320】

以上説明した本実施の形態においては、アクセスシーケンス抽出、検索キーワード抽出、質的評価値算出及び量的評価値算出算出に関する各処理が並列ではなく直列に実行されており、上記各実施の形態と同様の効果を得ることができる。また、本実施の形態においては、ログを一括ではなく1アクセスシーケンスごとに処理することも可能になる。

【0321】

なお、上記各実施の形態においては、一つのアクセスシーケンスに対して検索キーワードが一回入力された場合を想定して説明しているが、一つのアクセスシーケンスに対して検索キーワードが複数回入力された場合であっても検索キーワードの質的評価値、量的評価値を求めることができる。

【0322】

例えば、一つのアクセスシーケンスに対して検索キーワードが複数回入力されている場合、そのアクセスシーケンスを検索キーワードの入力を境界として複数のアクセスシーケンスに細分化し、この細分化されたアクセスシーケンスを用いて質的評価値、量的評価値を求めるとしてもよい。

【0323】

この場合、一つのアクセスシーケンスに対して検索キーワードが一回入力された場合における上記各実施の形態の処理方法を、細分化後の一つのアクセスシーケンスに対して適用することができる。

【0324】

(第7の実施の形態)

本実施の形態においては、分析対象のWebサイトのWebアクセスログ分析を実行し、分析対象ではない他のWebサイトに存在する外部コンテンツから分析対象のWebサイトが閲覧される傾向を分析するための検索キーワード分析システムについて説明する。

【0325】

本実施の形態では、外部コンテンツが検索システムであり、検索システムで入力された検索キーワードに基づいて閲覧の傾向を調査する。

【0326】

図13は、本実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図である。

【0327】

検索キーワード分析システム47は、記録媒体48に記録された検索キーワード分析プログラム49を読み込み、実行することにより、ログ前処理機能12a、アクセスシーケンス抽出機能13a、キーワード抽出機能14a、評価値算出

機能39a、カテゴリ形成機能52、トレンド分析機能53、出力処理機能54を実現する。また、検索キーワード分析システム47は、出力装置55、記録装置38、58～60を具備する。

【0328】

なお、ログ前処理機能12a、アクセスシーケンス抽出機能13a、キーワード抽出機能14a、評価値算出機能39a、カテゴリ形成機能52、トレンド分析機能53、出力処理機能54は、ハードウェアにより実現されてもよい。

【0329】

本実施の形態において、例えば、記録装置38への記録、読み出しを経由して各機能12a～14a、39a、52～54の間でデータの提供、受け付けが行われるとしてもよい。

【0330】

記録装置38、58～60としては、例えば内部メモリ、ハードディスクなどが適用される。各記録装置38、58～60は、データベースを構成する要素であり、各データはデータベースに管理されるとしてもよい。また、各記録装置38、58～60は、自由に統合可能である。

【0331】

記録装置4に記録されているWebアクセスログ7は、閲覧者によってアクセスされた部分を示すアクセス先識別データ、アクセスされた時間を示す時間データ（例えば日付、時刻）、閲覧者が検索システムによる検索結果に基づいてアクセスした場合の検索キーワードとを関連付けたデータを含む。

【0332】

ログ前処理機能12a、アクセスシーケンス抽出機能13a、キーワード抽出機能14a、評価値算出機能39aは、上述したログ前処理機能12、アクセスシーケンス抽出機能13、キーワード抽出機能14、評価値算出機能39と同様の処理を実行するが、Webアクセスログ7の時間データに基づいて、所定の期間毎のデータに対して処理を実行し、所定の期間毎に、分析対象のWebサイトをアクセスするために入力された各検索キーワードと質的評価値である満足度と量的評価値であるニーズ値とを対応付けた関係データを作成する点が異なる。

【0333】

なお、上記各機能12a～14a、39aによって、所定の期間毎に、関係データを求める処理を実行するためには、例えばログ前処理機能12aが、Webアクセスログ7のうち所定の期間毎のデータを受け付けることで可能となる。

【0334】

評価値算出機能39aは、所定の期間毎に、分析対象のWebサイトをアクセスするために入力された各検索キーワードの満足度とニーズ値とを求め、カテゴリ形成機能52とトレンド分析機能53に提供する。なお、例えば、同一期間内に、同一の検索キーワードの質的評価値又は量的評価値が複数算出された場合には、評価値算出機能39aは、その期間について、同一の検索キーワードの質的評価値又は量的評価値の平均を求めるとしてもよい。

【0335】

カテゴリ形成機能52は、記録装置58に記録されているカテゴリデータ72、所定の期間毎の各検索キーワードの満足度及びニーズ値に基づいて、各検索キーワードの属するカテゴリ単位で満足度及びニーズ値を求め、求めた各カテゴリの満足度及びニーズ値をトレンド分析機能53に提供する。

【0336】

カテゴリデータ72は、例えば、検索キーワードとこの検索キーワードの属するカテゴリとを関連付けたデータである。

【0337】

カテゴリ形成機能52は、例えば、期間毎に、カテゴリに属する検索キーワードのニーズ値の総和又は平均値を求め、期間毎のカテゴリのニーズ値とする。

【0338】

カテゴリ形成機能52は、例えば、期間毎に、カテゴリに属する各検索キーワードの満足度をニーズ値で重み付けした重み満足度の総和を、カテゴリに属する各検索キーワードのニーズ値の総和で割った値を求め、期間毎のカテゴリの満足度とする。

【0339】

トレンド分析機能53は、期間経過にそって、各検索キーワード及び各カ

リのニーズ値と満足度の遷移状況を表すトレンドデータ73を求め、記録装置59に記録する。

【0340】

トレンド分析機能53は、領域計算機能53a、領域判断機能53b、メッセージ付加機能53cを具備する。

【0341】

領域計算機能53aは、記録装置60に記録されているニーズ境界値74、満足度境界値75、評価値算出機能39aによって求められた期間毎の検索キーワードの満足度及びニーズ値、カテゴリ形成機能52によって求められた期間毎のカテゴリの満足度及びニーズ値に基づいて、期間毎に、検索キーワード及びカテゴリがニーズ境界値74、満足度境界値75によって決定されるどの領域に位置するか判断する。

【0342】

ニーズ境界値74、満足度境界値75は、ニーズ値、満足度の意味が切り換わる境界を表す値である。このニーズ境界値74、満足度境界値75によって、分析対象のWebサイトをアクセスするために入力された検索キーワード及びカテゴリの位置する領域が決定される。

【0343】

領域判断機能53bは、期間の切り換わりに応じて、すなわち期間の経過にそって検索キーワードが領域を移動したか、カテゴリが領域を移動したか判断する。

【0344】

メッセージ付加機能53cは、記録装置60に記録されているメッセージデータ76に基づいて、期間経過にそって領域を移動した検索キーワード又はカテゴリに、その移動した領域に対応するメッセージ（例えばコメント文、注釈など）を付する。

【0345】

メッセージデータ76は、例えば分析対象のWebサイトをアクセスするために入力された検索キーワードの移動前の領域と移動後の領域とメッセージとを関

連付けたデータである。なお、メッセージデータ76には、領域が変化した場合に、期間経過後の領域に移動したことを意味するメッセージを付するための規則が定義されていればよい。

【0346】

トレンド分析機能53は、期間毎の検索キーワード及びカテゴリのニーズ値及び満足度、期間毎の検索キーワード及びカテゴリの位置する領域、領域移動の有無を示す判断結果、領域を移動した検索キーワードに付されたメッセージ、領域を移動したカテゴリに付されたメッセージを含むトレンドデータ73を記録装置59に記録する。

【0347】

出力処理機能54は、記録装置59に記録されているトレンドデータ73を、出力装置55を用いて出力するための処理を実行する。出力処理機能54は、トレンドデータ73を例えばテーブル、グラフ、各種の図などの形式で出力する。

【0348】

図14は、本実施の形態に係る検索キーワード分析システム47の動作の一例を示すフローチャートである。

【0349】

ステップV1において、ログ前処理機能12a、アクセスシーケンス抽出機能13a、キーワード抽出機能14a、評価値算出機能39aは、Webアクセスログ7の時間データに基づいて、所定の期間毎に、分析対象のWebサイトをアクセスするために入力された各検索キーワードと満足度とニーズ値とを対応付けた関係データを作成する。

【0350】

ステップV2において、カテゴリ形成機能52cは、期間毎に、各検索キーワードの属するカテゴリのニーズ値、満足度を求める。

【0351】

ステップV3において、トレンド分析機能53の領域計算機能53aは、ニーズ境界値74、満足度境界値75に基づいて、期間毎に、検索キーワード及びカテゴリの位置する領域を求め、検索キーワード及びカテゴリのニーズ値及び満足

度、検索キーワード及びカテゴリの位置する領域をトレンドデータ73に含める。

【0352】

ステップV4において、トレンド分析機能53の領域判断機能53bは、期間経過にそって、検索キーワードの位置する領域が変化したか、カテゴリの位置する領域が変化したか判断し、判断結果をトレンドデータ73に含める。

【0353】

検索キーワード又はカテゴリの位置する領域が変化したと判断された場合、ステップV5において、トレンド分析機能53のメッセージ付加機能53cは、メッセージデータ76に基づいて、領域が変化した検索キーワード、領域が変化したカテゴリに、領域の変換に応じたメッセージを付し、結果をトレンドデータ73に含める。

【0354】

ステップV6において、出力処理機能54は、分析者の操作に応じて、トレンドデータ73に基づいて、検索キーワード、カテゴリのニーズ値、満足度の遷移状況とメッセージとを表示する。

【0355】

以下、評価値算出機能39aによって求められる関係データについて説明する。

【0356】

評価値算出機能39aは、期間毎に、図15に示すような、分析対象のWebサイトをアクセスするために入力された検索キーワードのニーズ値と満足度を計算した結果を示す関係データを、トレンド分析機能53の領域計算機能53a及びカテゴリ形成機能52に提供する。

【0357】

カテゴリ形成機能52は、期間毎に、図16に示すように、カテゴリデータ72と検索キーワードのニーズ値及び満足度とに基づいて、検索キーワード単位のニーズ値、満足度からカテゴリ単位のニーズ値、満足度を求め、カテゴリのニーズ値、満足度をトレンド分析機能53の領域計算機能53aに提供する。

【0358】

トレンド分析機能53は、図17に示すような検索キーワードの遷移状況にメッセージを付したデータ、及び図18に示すようなカテゴリの遷移状況にメッセージを付したデータをトレンドデータ73に含め、トレンドデータ73を記録装置59に記録する。

【0359】

以下に、カテゴリ形成機能52について詳細に説明する。

【0360】

図19は、カテゴリ形成機能52の動作の一例を示すフローチャートである。

【0361】

ステップW1において、カテゴリ形成機能52は、カテゴリデータ72に基づいて、期間毎に、あるカテゴリに属する検索キーワードを識別する。

【0362】

ステップW2において、カテゴリ形成機能52は、期間毎に、(1)式に示すようにカテゴリに属する各検索キーワードのニーズ値の総和を求め、カテゴリのニーズ値とする。

【0363】**【数1】**

$$\text{カテゴリのニーズ値} = \sum (\text{カテゴリの } i \text{ 番目の検索キーワードのニーズ値}) \cdots (1)$$

【0364】

ステップW3において、カテゴリ形成機能52は、期間毎に、(2)式に示すように、カテゴリに属する各検索キーワードの満足度の加重平均を求め、カテゴリの満足度とする。

【0365】

【数2】

$$\text{カテゴリの満足度} = \frac{\sum_{i=1}^n \{ \text{カテゴリの } i \text{ 番目の検索キーワードのニーズ値} \} \times \{ \text{カテゴリの } i \text{ 番目の検索キーワードの満足度} \}}{\sum_{i=1}^n \{ \text{カテゴリの } i \text{ 番目の検索キーワードのニーズ値} \}}$$

… (2)

【0366】

この数式（2）では、カテゴリの満足度は、そのカテゴリに属する各検索キー

ワードの満足度とニーズ値との積の総和を、そのカテゴリに属する検索キーワードのニーズ値の総和で割って求められる。

【0367】

ステップW4において、カテゴリ形成機能52は、期間毎に、全てのカテゴリについてニーズ値及び満足度を求めたか判断し、全てのカテゴリについてニーズ値及び満足度を求めていない場合、上記ステップW1以下の処理を繰り返す。

【0368】

以下に、カテゴリの満足度を、そのカテゴリに属する各検索キーワードの満足度をそれぞれのニーズ値で重み付けした値の総和を、そのカテゴリに属する各検索キーワードのニーズ値の総和で割って求める理由について説明する。

【0369】

例えば、あるカテゴリCにおいて、検索キーワードC₁のニーズ値は極めて小さいが、検索キーワードC₁の満足度は極めて大きいとする。この場合、検索キーワードC₁を用いて分析対象のWebサイトにアクセスした閲覧者数は極めて少ないが、この小人数の閲覧者だけは満足していると判断される。

【0370】

一方、同一カテゴリCにおいて、検索キーワードC₂のニーズ値は極めて大きいが、検索キーワードC₂の満足度は極めて小さいとする。この場合、検索キーワードC₂を用いて分析対象のWebサイトにアクセスした閲覧者数は極めて多いが、この多人数の閲覧者は満足していないと判断される。

【0371】

このような場合に、単にカテゴリCの満足度を、このカテゴリCに属する検索キーワードC₁、C₂の満足度の平均とすると、閲覧者数の少なかった検索キーワードC₁の満足度と閲覧者の多かった検索キーワードC₂の満足度とが同レベルに扱われる。したがって、満足していない閲覧者が大多数であるにもかかわらず、カテゴリは中程度の満足度と判断される。

【0372】

しかしながら、カテゴリCの満足度は、閲覧者数の少なかった検索キーワードC₁の満足度よりも閲覧者の数が多い検索キーワードC₂の満足度から影響を受け

るようすべきである。

【0373】

そこで、カテゴリCに属する各検索キーワードC₁、C₂の満足度をニーズ値で重み付けし、合計を求める。

【0374】

そして、このカテゴリCと他のカテゴリとの間で、それぞれの閲覧者数に関係なく等しい基準で満足度を評価するために、カテゴリCについて、重み付けした満足度の合計を、そのカテゴリCの全ての閲覧者数すなわちそのカテゴリCに属する検索キーワードのニーズ値の合計で割る。

【0375】

図20は、カテゴリ形成機能52によるデータの変換結果の一例を示す図である。

【0376】

カテゴリ形成機能52は、検索キーワードとその検索キーワードの属するカテゴリとを関連付けたカテゴリデータ72と、期間毎の検索キーワードのニーズ値と満足度を示す関係データ77を受け付ける。

【0377】

そして、カテゴリ形成機能52は、期間毎に、カテゴリのニーズ値と満足度とを求めたカテゴリライズデータ78を求める。この図20の例では、5種類の検索キーワードが2種類のカテゴリに集約されている。各カテゴリについては、個々の検索キーワードに対する分析と同様の分析を行うことができる。

【0378】

以下に、トレンド分析機能53について詳細に説明する。

【0379】

図21は、トレンド分析機能53の動作の一例を示すフローチャートである。

【0380】

ステップX1において、トレンド分析機能53の領域計算機能53aは、期間毎の全ての検索キーワード及びカテゴリについてそれぞれのニーズ値及び満足度と、ニーズ境界値74及び満足度境界値75とに基づいて、期間毎の全ての検索

キーワード及びカテゴリの位置する領域を求め、検索キーワード及びカテゴリのニーズ値及び満足度、領域をトレンドデータ73に含める。

【0381】

ステップX2において、トレンド分析機能53の領域判断機能53bは、期間経過の前後で、検索キーワード及びカテゴリの位置する領域を比較し、検索キーワード及びカテゴリの領域が変化するか判断し、領域の変化する検索キーワード又はカテゴリを求め、結果をトレンドデータ73に含める。

【0382】

検索キーワード又はカテゴリの位置する領域が変化する場合、ステップX3において、トレンド分析機能53のメッセージ付加機能53cは、メッセージデータ76に基づいて、領域の変化する検索キーワード又はカテゴリに対して、期間経過前と後の領域に対応するメッセージを関連付け、トレンドデータ73に含める。

【0383】

図22は、ニーズ境界値74と満足度境界値75とによって決定される領域の名称の一例を示す図である。

【0384】

ニーズ値についてニーズ境界値74、満足度について満足度境界値75が設定されており、平面が4分割され、4つの領域に名称が付されている。

【0385】

この例では、ニーズ境界値74及び満足度境界値75をともに100としている。そして、ニーズ値と満足度の小さい領域を「失敗領域」、ニーズ値は小さいが満足度は大きい領域を「隠れ成功領域」、ニーズ値は大きいが満足度は小さい領域を「チャンスロス領域」、ニーズ値と満足度の大きい領域を「成功領域」としている。

【0386】

図23は、トレンド分析機能53によって作成されるトレンドデータ73の一例を示す図である。

【0387】

例えば、期間経過の切り換わり前と後で領域が変化した検索キーワード又はカテゴリには、移動先の領域に遷移したことを意味するメッセージが付される。この23のトレンドデータでは、領域変動のあった検索キーワード「食べ放題」「ランチ」にそれぞれメッセージ「成功領域へ遷移」「失敗領域へ遷移」というメッセージが付されている。

【0388】

このようなメッセージの付与は、期間経過後の領域とその領域に対応したメッセージとを関連付けたメッセージデータ76を用意し、トレンド分析機能53の領域判断機能53bによって期間経過前後で領域の変化した検索キーワード又はカテゴリを検出し、トレンド分析機能53のメッセージ付加機能53cによって、メッセージデータ76に基づいて、領域が変化した検索キーワード又はカテゴリに期間経過後に位置する領域に対応するメッセージを付することで可能である。

【0389】

図24は、領域が移動した場合に付与されるメッセージの他の例を示す図である。

【0390】

期間経過前後で境界を超えて変化する検索キーワード又はカテゴリに、どの境界をどの方向に超えたのか表すメッセージが付される。例えば、「隠れ成功領域」から「成功領域」に移動した検索キーワード又はカテゴリには、「ニーズ上昇」というメッセージを付する。他の場合も同様である。

【0391】

このようなメッセージの付与は、期間経過前の領域、期間経過後の領域、メッセージとを関連付けたメッセージデータ76を用意し、トレンド分析機能53の領域判断機能53bによって期間経過前後に領域の変化した検索キーワード又はカテゴリを検出し、トレンド分析機能53のメッセージ付加機能53cによって、メッセージデータ76に基づいて、領域の変化した検索キーワード又はカテゴリの期間経過前の領域と期間経過後に位置する領域に関連付けされたメッセージを求めることが可能である。

【0392】

以下に、出力処理機能54について詳細に説明する。

【0393】

出力処理機能54は、例えば図25に示すようなグラフを表示する。

【0394】

このグラフは、ニーズ値と満足度を表す座標軸を持ち、ニーズ境界値と満足度境界値とを表示し、各領域に領域名称を記述し、期間毎の検索キーワード又はカテゴリを示す指標をニーズ値と満足度に基づいてプロットし、プロットされた指標が線で結ばれて描かれている。なお、この図25では、期間経過毎の遷移状況を示す軌跡が矢印で表されている。

【0395】

その他にも、出力処理機能54は、例えば上記図17、18に示すようなニーズ値及び満足度に基づく期間毎の検索キーワード又はカテゴリのグラフ、上記図20に示すような期間毎のカテゴリ化データ78、上記図23に示すような検索キーワードの遷移状況とメッセージとを表すテーブル形式のトレンドデータ73を出力装置55に表示する。

【0396】

以上説明した本実施の形態においては、分析者は、分析対象のWebサイトについて集客戦略、コンテンツ改善の指針を得ることができる。

【0397】

例えば分析者は以下の効果を得ることができる。

【0398】

第1に、分析者は、閲覧者数が多いが閲覧者がすぐに分析対象のWebサイトの閲覧を終了するような検索キーワード及びカテゴリを調査でき、検索キーワード及びカテゴリに基づいて検索された分析対象のWebサイトが閲覧者の要求を満たすか否か判断できる。また、分析者は、閲覧者の要求を満たすことができないという一種のチャンスロスの状態を検出し、調査できる。したがって、分析者は、分析対象のWebサイトにおけるチャンスロスを改善し、分析対象のWebサイトの質を向上させることができる。

【0399】

第2に、分析者は、分析対象のWebサイトをアクセスする閲覧者の数を増加させ、かつアクセスした閲覧者を満足させる良好な検索キーワード及びカテゴリが存在するか判断できる。さらに、分析者は、このような検索キーワード及びカテゴリの良好な状態が一時的又は減少傾向にあるか判断でき、新たな検索キーワード及びカテゴリを検討すべきか判断できる。

【0400】

第3に、分析者は、個別の検索キーワード単位のみではなく、各検索キーワードの属するカテゴリ単位で静的又は動的な分析を行うことができる。このため、分析者は、自己の検討したい内容、検討したい事項に合わせて分析結果を得ることができ、分析対象のWebサイトについての閲覧傾向を適切かつ容易に把握でき、分析の精度を向上させることができる。

【0401】

第4に、分析者は、検索キーワード及びカテゴリの分析結果の時間的推移をグラフにより得ることができ、閲覧の傾向の時間的変化を容易に把握でき、分析作業を効率化できる。

【0402】

第5に、分析者は、検索キーワード及びカテゴリのトレンド分析の結果を用いて、膨大なデータから有用なデータを迅速に得ることができ、分析の精度を向上させることができ、分析に必要な労力を削減できる。

【0403】

第6に、分析者は、従来のWebアクセスログによる単純な集計ではなく、閲覧者の流行を容易かつ高精度に調査することができる。

【0404】

第7に、分析者は、検索キーワードを個別に分析しても把握できないようなカテゴリ全体のトレンドを把握できる。

【0405】

第8に、分析者は、閲覧者の意思で様々な形式で入力されつ検索キーワードについて容易かつ適切に自らの意思にそった分析結果を得ることができる。以下に

、この第8の効果について具体例を挙げて説明する。

【0406】

検索キーワードは、閲覧者の意思で様々な形式で入力されるため、Webアクセスログ分析により検索キーワードについて分析しても、分析者は自らの意思にそった分析結果を得ることが困難な場合がある。

【0407】

例えば、分析者は、料理関係の検索キーワード入力後に分析対象のWebサイトがどのようなアクセス状態か調べたいとする。この場合、従来のWebアクセスログ分析では、検索キーワード「料理」について閲覧者数を分析結果として求めることはできるが、料理と関係する他の検索キーワード「グルメ」「ファーストフード」などについての閲覧者数まで分析結果に含めることはできない。

【0408】

しかしながら、本実施の形態のようにカテゴリ単位で分析することによって、分析者は、自己の求める分析結果を適切に取得することができる。

【0409】

(第8の実施の形態)

本実施の形態においては、上記第7の実施の形態の変形例について説明する。本実施の形態に係る検索キーワード分析システムは、上記第7の実施の形態とトレンド分析機能が主に相違し、他の部分は同様であるとする。

【0410】

本実施の形態に係るトレンド分析機能は、ニーズ値境界値と満足度境界値とを用いることなく、ニーズ値と満足度を座標軸とする平面上における期間経過前後の検索キーワード及びカテゴリの移動距離と移動方向を求める。

【0411】

図26は、本実施の形態に係るトレンド分析機能の構成の一例を示すブロック図である。

【0412】

本実施の形態に係るトレンド分析機能79は、期間毎の検索キーワード及びカテゴリのニーズ値、満足度を座標値とし、期間経過前後における全ての検索キ

ワード及びカテゴリの移動距離を求め、移動距離の大きい検索キーワード及びカテゴリについて、その方向に応じたメッセージを付す。

【0413】

このトレンド分析機能79は、移動距離計算機能79a、移動距離判断機能79b、移動方向計算機能79c、メッセージ付加機能79dとを具備する。なお、トレンド分析機能79は、上記第7の実施の形態で説明したトレンド分析機能53と同じ機能も具備するとする。

【0414】

移動距離計算機能79aは、評価値算出機能39aによって求められた期間毎の検索キーワード、カテゴリ形成機能52によって求められた期間毎のカテゴリのニーズ値、満足度に基づいて、式（3）にしたがって、期間経過前後における全ての検索キーワード及びカテゴリの移動距離を求め、結果を移動距離判断機能79bに提供すると共にトレンドデータ81に含める。

【0415】

【数3】

… (3)

$$\text{移動距離} = \sqrt{\{(\text{後のニーズ値}) - (\text{前のニーズ値})\}^2 + \{(\text{後の満足度}) - (\text{前の満足度})\}^2}$$

【0416】

移動距離判断機能79bは、記録装置60に記録されている移動距離しきい値80を超えて期間毎の検索キーワード又はカテゴリが移動したか判断すると共に

、結果をトレンドデータ81に含める。

【0417】

表24に、移動距離しきい値の具体例とトレンド分析機能79の動作の関係を示す。この表24の例では、移動距離しきい値が120と設定されており、移動距離が120以上の場合に、移動方向の計算及びメッセージ付与が行われることを示している。

【0418】

【表24】

(移動距離しきい値120の場合)

移動距離D	行なうべき動作
$D < 120$	—
$120 \leq D$	移動方向の計算、メッセージの付与

【0419】

移動方向計算機能79cは、移動距離が移動距離しきい値を超えた検索キーワード又はカテゴリのニーズ値、満足度に基づいて、式(4)にしたがって、期間経過前後の移動方向を計算し、結果をメッセージ付加機能79dに提供すると共にトレンドデータ81に含める。

【0420】

【数4】

… (4)

$$\text{移動方向} = \tan^{-1} \frac{(\text{後の満足度}) - (\text{前の満足度})}{(\text{後のニーズ値}) - (\text{前のニーズ値})}$$

【0421】

メッセージ付加機能79dは、記録装置60に記録されているメッセージデータ82に基づいて、移動距離しきい値を超えた検索キーワード又はカテゴリに対

して移動方向に応じたメッセージを付する。

【0422】

表25に、メッセージデータ82の一例を示す。移動方向を表す角度と、ニーズ及び満足の上昇、下降の意味するメッセージが対応付けられている。

【0423】

【表25】

移動方向 θ	付加するメッセージ
$0^\circ \leq \theta < 30^\circ$	ニーズ上昇
$30^\circ \leq \theta < 60^\circ$	ニーズ・満足度とも上昇
$60^\circ \leq \theta < 120^\circ$	満足度上昇
$120^\circ \leq \theta < 150^\circ$	ニーズ下降・満足度上昇
$150^\circ \leq \theta < 210^\circ$	ニーズ下降
$210^\circ \leq \theta < 240^\circ$	ニーズ・満足度とも下降
$240^\circ \leq \theta < 300^\circ$	満足度下降
$300^\circ \leq \theta < 330^\circ$	ニーズ上昇・満足度下降
$330^\circ \leq \theta < 360^\circ$	ニーズ上昇

【0424】

図27は、メッセージと移動方向の関係の具体例を示す。この図27において、移動方向 0° 以上～ 30° 未満の場合にはメッセージ「ニーズ上昇」と表示すると定義され、移動方向 30° 以上～ 60° 未満の場合にはメッセージ「ニーズ・満足度とも上昇」と表示すると定義され、移動方向 60° 以上～ 90° 未満の場合には「満足度上昇」と表示すると定義されている。

【0425】

なお、ニーズ値と満足度の双方が増加する移動方向 0° 以上～ 90° 未満の範囲のみではなく、ニーズ値と満足度の少なくとも一方が減少する 90° 以上～ 360° 未満範囲のメッセージも同様に定義できる。

【0426】

図28は、トレンド分析機能79によって作成されたトレンドデータ81の一例を示す図である。

【0427】

移動距離の大きい検索キーワード「食べ放題」「ランチ」に、移動方向に対応するメッセージが付されている。

【0428】

図29は、本実施の形態に係る検索キーワード分析システムの動作の一例を示すフローチャートである。

【0429】

ステップY1、Y2は、上記図3のステップV1、V2と同様であるため、説明を省略する。

【0430】

ステップY3において、トレンド分析機能79の移動距離計算機能79aは、期間経過前後における検索キーワード及びカテゴリの移動距離を求め、トレンドデータ81に含めるとともに移動距離判断機能79bに提供する。

【0431】

ステップY4において、トレンド分析機能79の移動距離判断機能79bは、記録装置60に記録されている移動距離しきい値80に基づいて、期間経過前後で検索キーワード及びカテゴリが移動距離しきい値80を超えるか判断し、結果をトレンドデータ81に含めるとともに移動方向計算機能79cに提供する。

【0432】

ステップY5において、トレンド分析機能79の移動方向計算機能79cは、移動距離が移動距離しきい値を超えたと判断された検索キーワード又はカテゴリの移動方向を計算し、結果をトレンドデータ81に含めるとともにメッセージ付加機能79dに提供する。

【0433】

ステップY6において、トレンド分析機能79のメッセージ付加機能79dは、記録装置60に記録されているメッセージデータ82に基づいて、移動方向の計算された検索キーワード又はカテゴリについてその移動方向に対応するメッセージを付加し、結果をトレンドデータ81に含める。

【0434】

ステップY7において、出力処理機能54は、分析者の操作に応じて、トレン

トデータ81に基づいて、検索キーワード又はカテゴリのニーズ値及び満足度の遷移状況とメッセージとを表示する。

【0435】

図30は、本実施の形態に係る出力処理機能54によって出力されたトレンドデータ81の一例を示す図である。

【0436】

この図30では、3期間以上でトレンド分析を行った結果得られたトレンドデータ81が表示されている。出力処理機能54は、連続する2期間毎に、トレンド分析機能79によって作成されたトレンドデータ81を同一平面上に表示する。この図30では、上記第1の実施の形態で説明したトレンドデータ73と本実施の形態で説明したトレンドデータ81とが組み合わされている。

【0437】

すなわち、この図30に示すグラフは、ニーズ値と満足度を表す座標軸を持ち、ニーズ境界値と満足度境界値とを表示し、各領域に領域名称を記述し、複数の期間毎の検索キーワード「宴会」を示す指標をそれぞれのニーズ値と満足度に基づいてプロットし、プロットされた指標が線で結ばれて描かれている。なお、この図30では、期間経過毎の遷移状況を示す軌跡が矢印で表されている。そして、図30では、領域の境界を超えた場合、及び移動距離の大きい場合のいずれかが成立した変化に対してメッセージが表示されている。

【0438】

以上説明した本実施の形態では、期間経過毎に、分析対象のWebサイトをアクセスするために入力される検索キーワードを分析でき、その変化量を移動距離、変化内容を移動方向として分析者に提供することができる。

【0439】

なお、上記第7及び本実施の形態においては、分析対象のWebサイトをアクセスするために入力される検索キーワードについて分析を行う場合について説明している。しかしながら、分析対象のWebサイトをアクセスする前にアクセスしていた外部コンテンツの存在する参照元のWebサイト又はドメインについて分析を行う場合にも同様に適用できる。

【0440】

その他にも、分析対象のWebサイトがキャンペーンメールに基づいてアクセスされた場合にこのキャンペーンメールについて分析を行い、キャンペーンの種別を表すIDについて分析結果を求めてよい。

【0441】

これにより、分析者は、検索キーワード、参照元のWebサイト、キャンペーンについてニーズの程度及び満足の程度を分析し、例えば空間上に表して結果を把握できる。

【0442】

また、上記第7及び本実施の形態の発明において、期間毎の関係データを求めるまでの処理は、上述した他の処理と変更可能である。すなわち、期間毎の関係データを求めるまでの処理として、上記各実施の形態で説明した各種手法を所定期間毎に実行することにより、同様の効果が得られる。

【0443】

また、上記第7及び本実施の形態の発明においては、カテゴリのニーズ値及び満足度を上記式（1）、式（2）で求めるとしているが、他の定義により他の手法でカテゴリのニーズ値及び満足度を求めるとしてもよい。

【0444】

また、本実施の実施の形態においては、検索キーワード及びカテゴリの移動距離及び移動方向を上記式（3）、式（4）で求めるとしているが、他の定義により他の手法で検索キーワード及びカテゴリの移動距離及び移動方向を求めるとしてもよい。

【0445】

また、上記第7及び本実施の形態においては、分析の対象はWebサイト単位でなくともよく、Webページ単位、複数のWebページから構成される任意のWebページグループ単位であってもよい。

【0446】

また、本実施の形態においては、移動距離が移動距離しきい値を超えた検索キーワード又はカテゴリのみ移動方向を算出するとしているが、全ての検索キーワード

ード又はカテゴリについて移動距離と移動方向を算出するとしてもよく、他の条件を満たす検索キーワード又はカテゴリについて移動距離と移動方向を算出するとしてもよい。

【0447】

また、上記第7及び本実施の形態発明においては、分析対象の閲覧される傾向の分析視点を、閲覧者が分析対象のWebサイトに期待する度合いを表すニーズ値と、分析対象のWebサイトが閲覧者に対して与える満足の度合いを表す満足度としている。しかしながら、閲覧の傾向の分析視点は、自由に変更できる。例えば、分析対象のWebサイトの目的が達成される機会の数と、目的が達成された数を表す達成度とを分析視点としてもよい。例えば、注文受付を目的とするWebサイトにおいて閲覧者の数を目標達成の機会数とし、注文を受け付けた数を達成度とし、この機会数と達成度を分析視点としてもよい。

【0448】

上記各実施の形態に係る検索キーワード分析システムの各構成要素は、同様の作用を実現可能であれば配置を変更させてもよく、また各構成要素を自由に組み合わせてもよい。

【0449】

また、上記各実施の形態において、検索キーワード分析システムは複数の計算機により構成され、検索キーワード分析プログラムは複数の計算機に分散して配置され、互いに連携を取りつつ処理を実行するとしてもよい。

【0450】

上記各実施の形態に係る検索キーワード分析プログラムは、例えば磁気ディスク（フレキシブルディスク、ハードディスク等）、光ディスク（CD-ROM、DVD等）、半導体メモリなどの記録媒体に書き込んでコンピュータに適用可能である。また検索キーワード分析プログラムは、通信媒体により伝送してコンピュータに適用することも可能である。コンピュータは、記録媒体に記録された検索キーワード分析プログラムを読み込み、検索キーワード分析プログラムによって動作が制御されることにより、上記の機能を実現する。

【0451】

【発明の効果】

以上詳記したように本発明においては、閲覧者の検索結果に対する質的評価値を迅速かつ高精度に求めることができ、この質的評価値に基づいてWebページ又はWebサイトのデザインやコンテンツ、検索システムを改善することができる。

【図面の簡単な説明】

【図1】 本発明の第1の実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図。

【図2】 同実施の形態に係る顧客サイトの構成の一例を示すブロック図。

【図3】 同実施の形態に係る検索キーワード分析システムの動作の一例を示すフローチャート。

【図4】 同実施の形態に係るキーワード抽出機能の動作の一例を示すフローチャート。

【図5】 同実施の形態に係る満足度算出機能の動作の一例を示すフローチャート。

【図6】 本発明の第2の実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図。

【図7】 本発明の第3の実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図。

【図8】 同実施の形態において表示されるグラフの一例を示す図。

【図9】 本発明の第4の実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図。

【図10】 本発明の第5の実施の形態に係る顧客サイトの構成の一例を示す図。

【図11】 サイト外検索システムから顧客サイトへの閲覧遷移の一例を示すブロック図。

【図12】 本発明の第6の実施の形態に係る顧客サイトの構成の一例を示す図。

【図13】 本発明の第7の実施の形態に係る検索キーワード分析システム

の構成の一例を示すブロック図。

【図14】 同実施の形態に係る検索キーワード分析システムの動作の一例を示すフローチャート。

【図15】 分析対象のWebサイトをアクセスするために入力された検索キーワードのニーズ値と満足度を計算した結果の一例を示す図。

【図16】 検索キーワード単位のニーズ値と満足度とカテゴリ単位のニーズ値と満足度の関係の一例を示す図。

【図17】 検索キーワードの遷移状況にメッセージを付した状態の一例を示す図。

【図18】 カテゴリの遷移状況にメッセージを付した状態の一例を示す図

。

【図19】 カテゴリ形成機能の動作の一例を示すフローチャート。

【図20】 カテゴリ形成機能によるデータの変換結果の一例を示す図

【図21】 同実施の形態に係るトレンド分析機能の動作の一例を示すフローチャート。

【図22】 ニーズ境界値と満足度境界値とによって決定される領域の名称の一例を示す図。

【図23】 同実施の形態に係るトレンド分析機能によって作成されるトレンドデータの一例を示す図

【図24】 同実施の形態において領域が移動した場合に付与されるメッセージの他の例を示す図。

【図25】 同実施の形態に係る出力処理機能によって表示されるグラフの一例を示す図。

【図26】 本発明の第8の実施の形態に係るトレンド分析機能の構成の一例を示すブロック図。

【図27】 メッセージと移動方向の関係の具体例を示す図。

【図28】 同実施の形態に係るトレンド分析機能によって作成されたトレンドデータの一例を示す図。

【図29】 同実施の形態に係る検索キーワード分析システムの動作の一例

を示すフローチャート。

【図30】 同実施の形態に係る出力処理機能によって出力されたトレンドデータの一例を示す図。

【図31】 検索システムを用いた代表的なWebページの遷移の例を示す図。

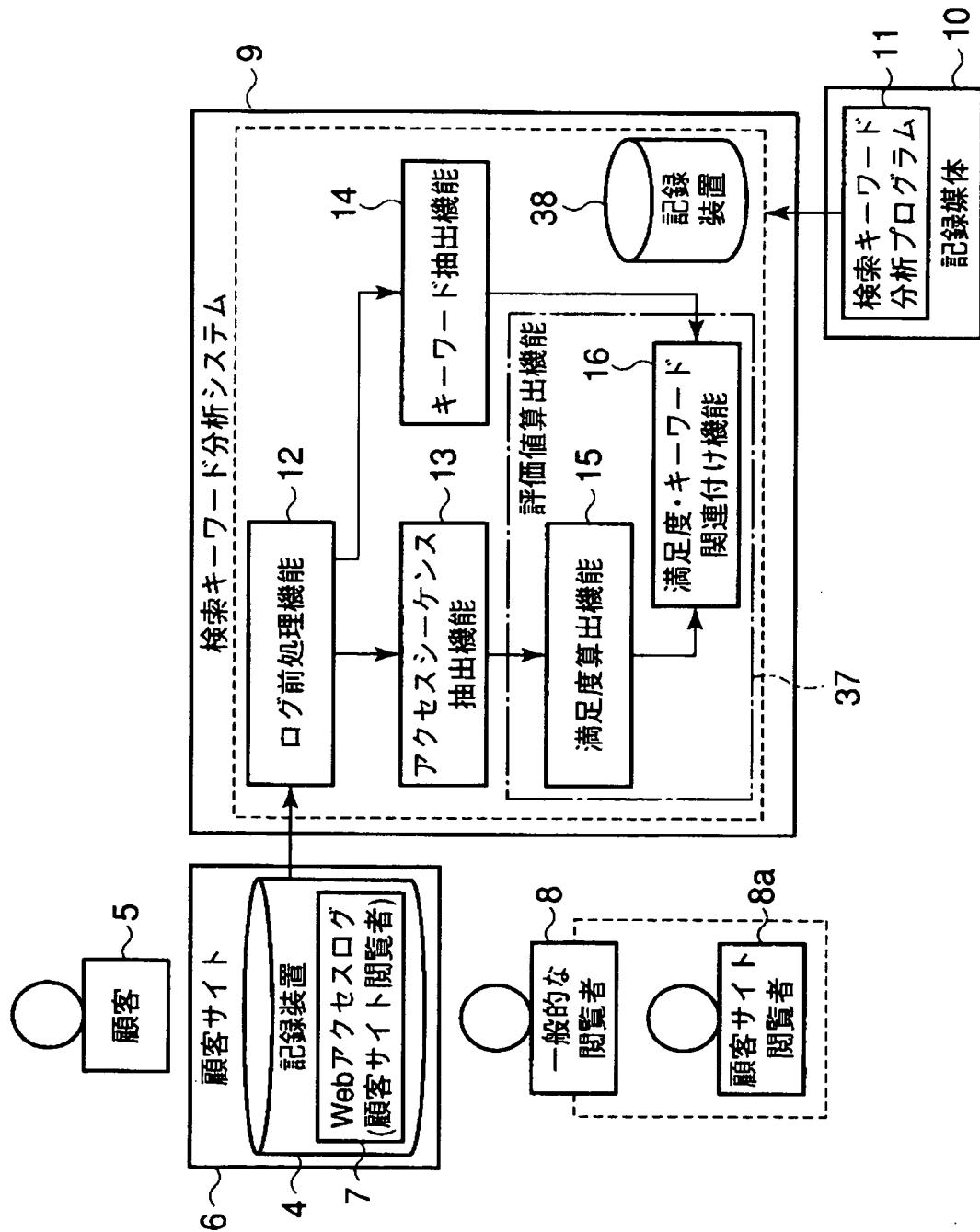
【符号の説明】

4、38…記録装置、5…顧客、6…顧客サイト、6a…サイト内検索システム、7…Webアクセスログ、8…一般的な閲覧者、8a…顧客サイト閲覧者、9, 18, 24, 29, 42…検索キーワード分析システム、10, 19, 25, 30, 43…記録媒体、11, 20, 26, 31, 44…検索キーワード分析プログラム、12, 12a…ログ前処理機能、13, 13a…アクセスシーケンス抽出機能、14, 14a, 45…キーワード抽出機能、15, 15a…満足度算出機能、16…満足度・キーワード関連付け機能、17, 23, 28, 33…出力データ、21, 21a…ニーズ値算出機能、22, 22a…満足度・ニーズ値関連付け機能、27…グラフ化機能、32…成功度算出機能、341～34n…サイト外検索システム、37, 39, 39a, 40, 46…評価値算出機能、41…成功度・キーワード関連付け機能、52…カテゴリ形成機能、53, 79…トレンド分析機能、54…出力処理機能

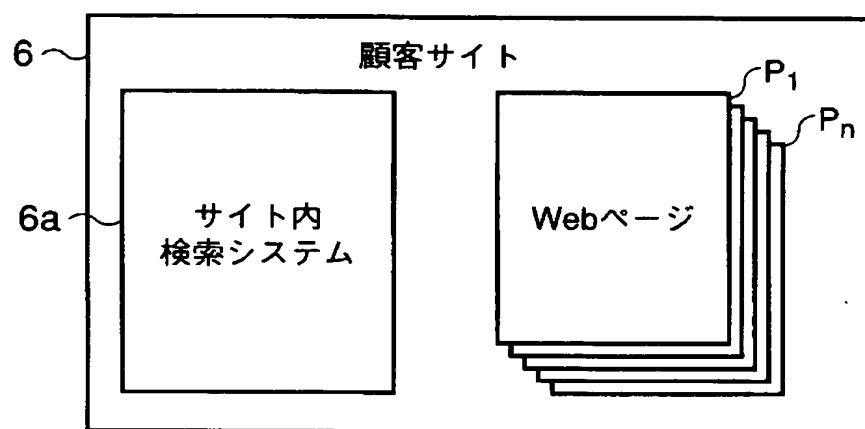
【書類名】

四面

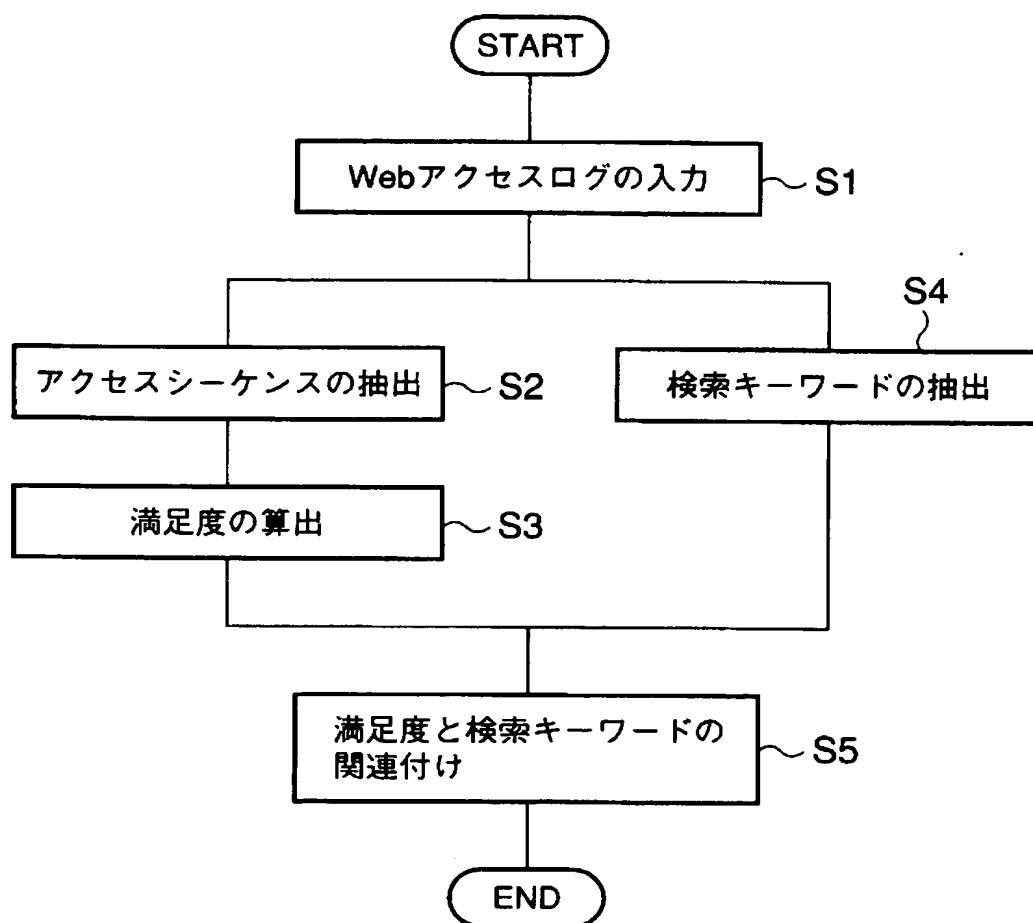
【圖 1】



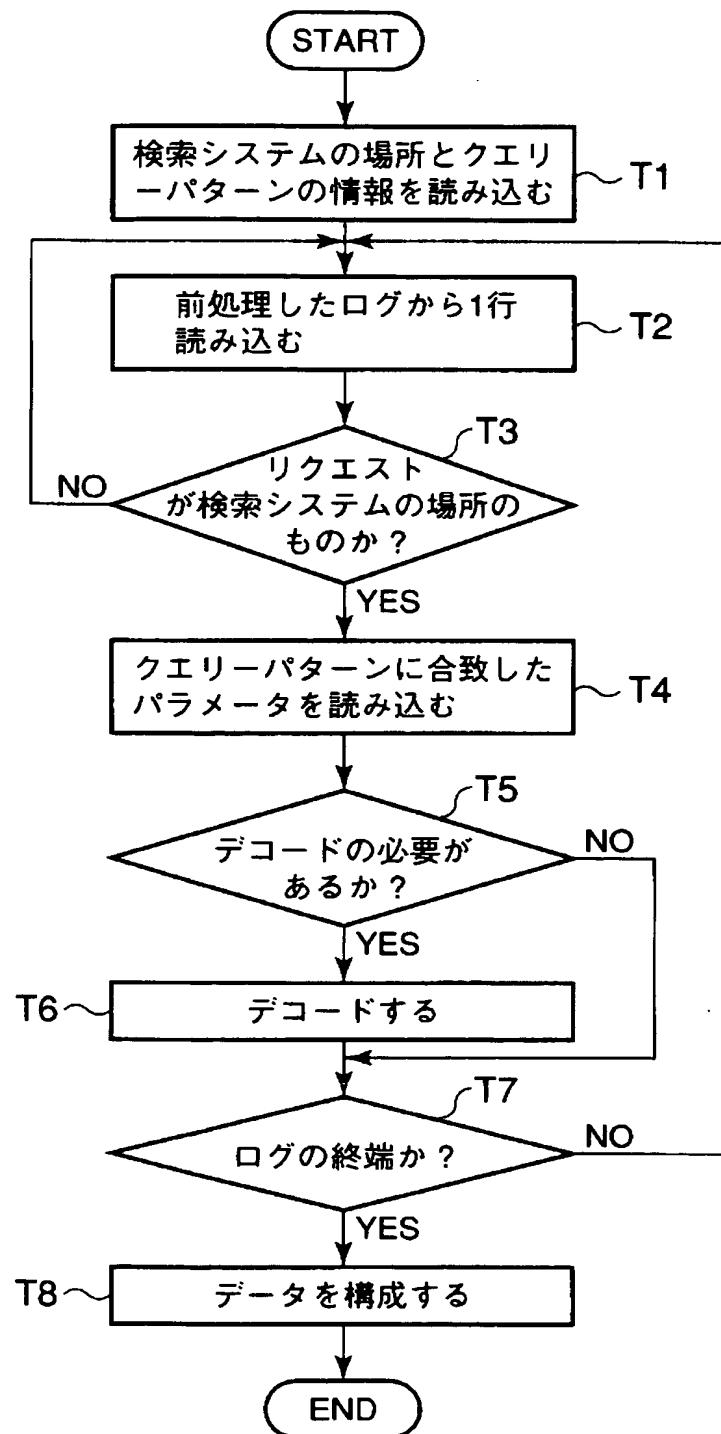
【図2】



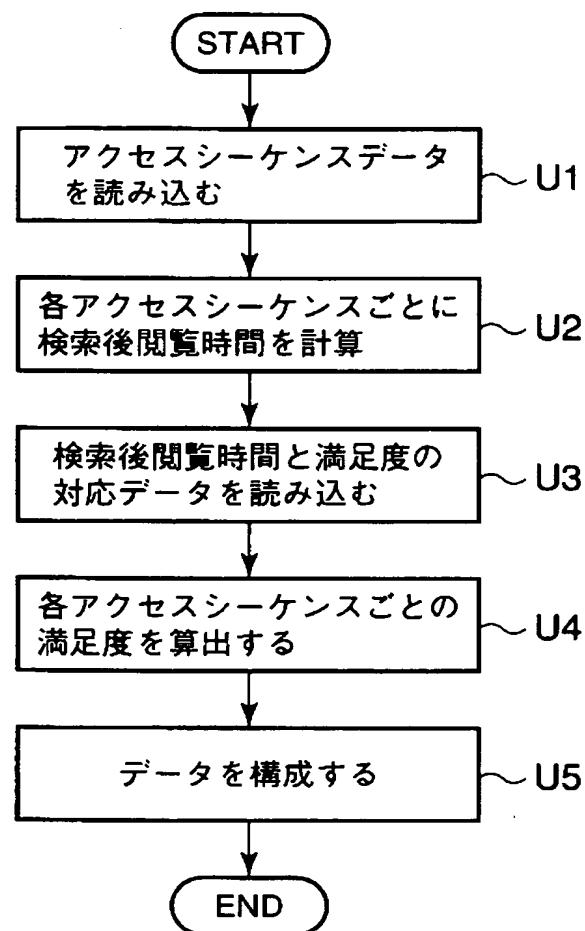
【図3】



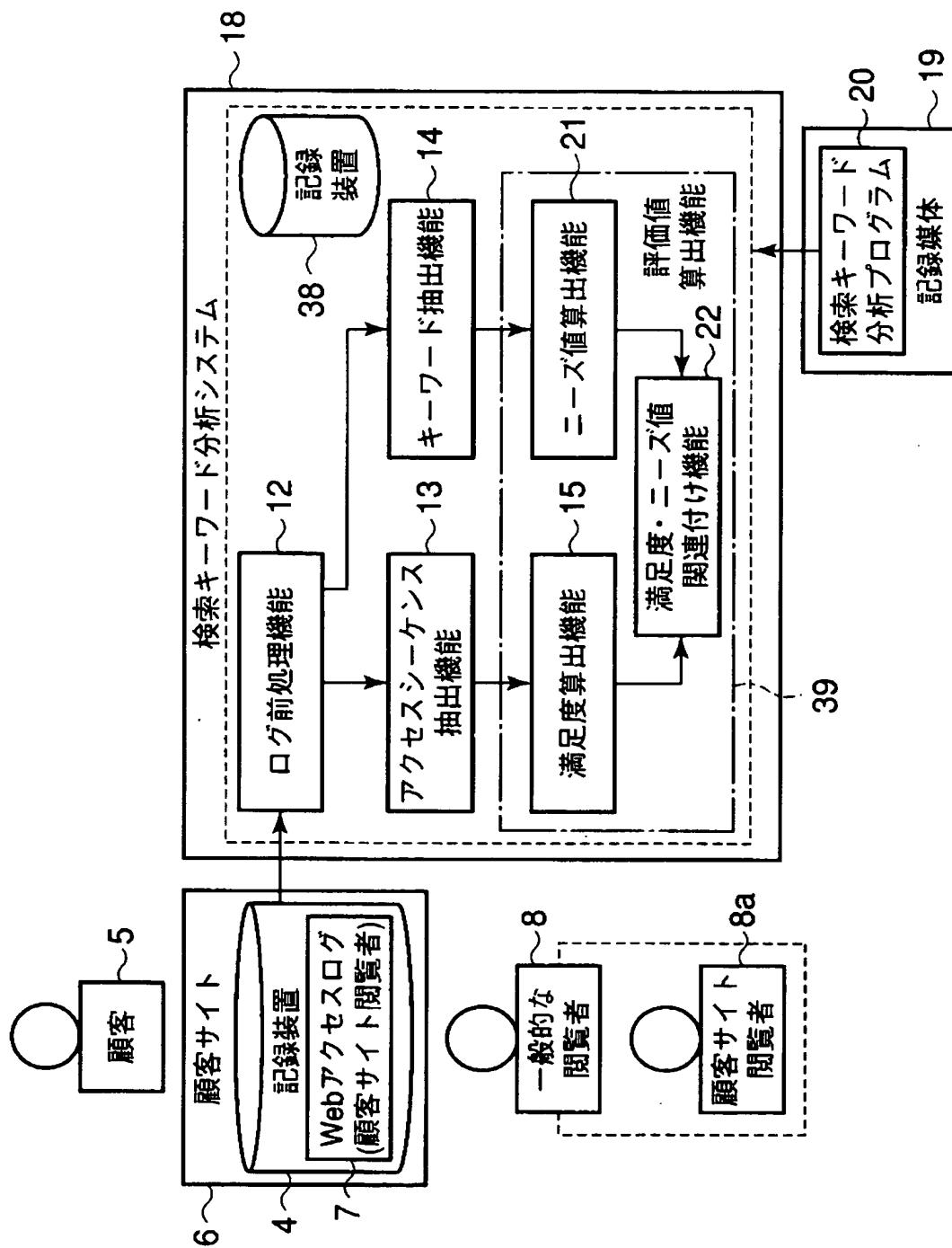
【図4】



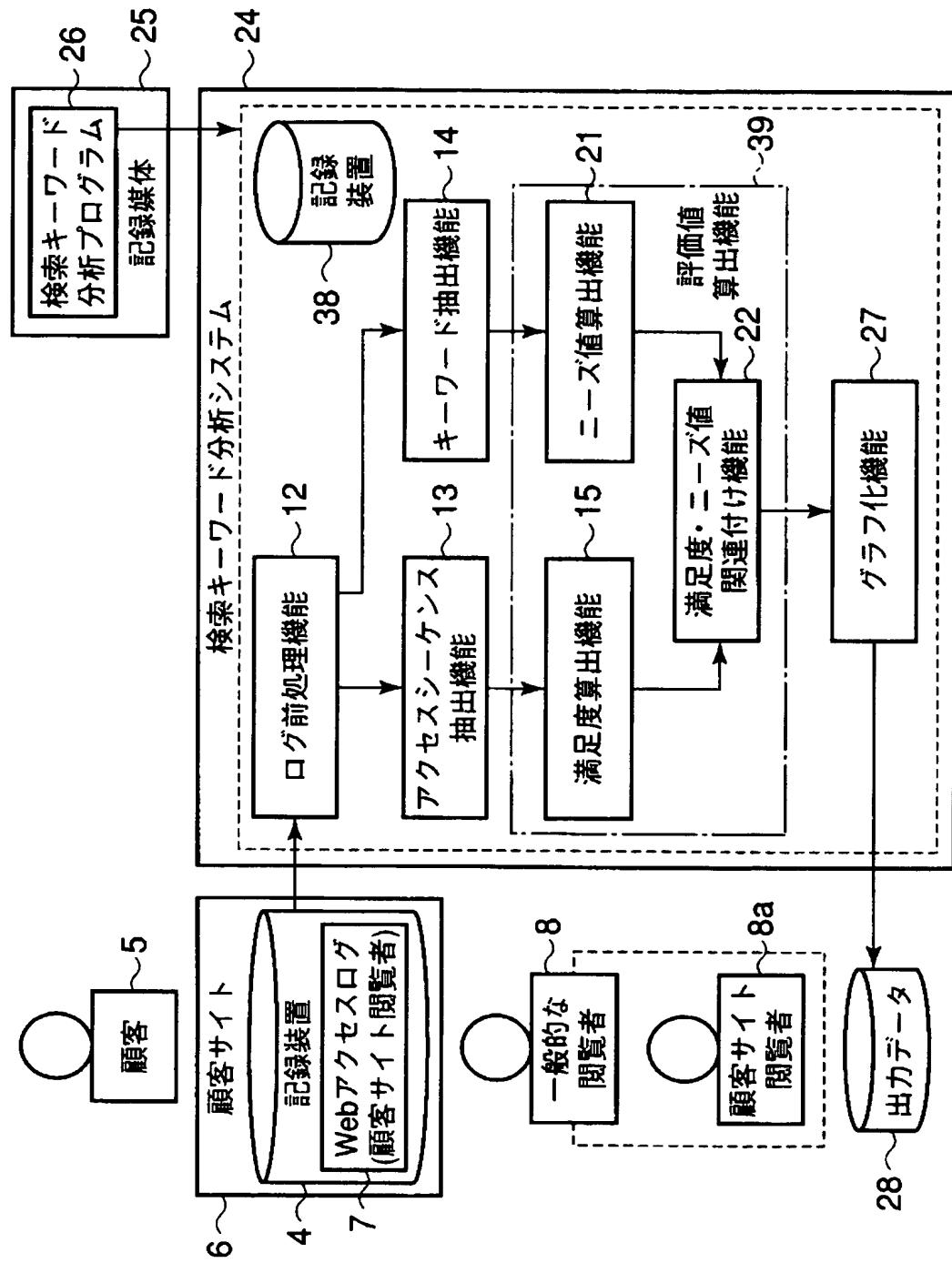
【図 5】



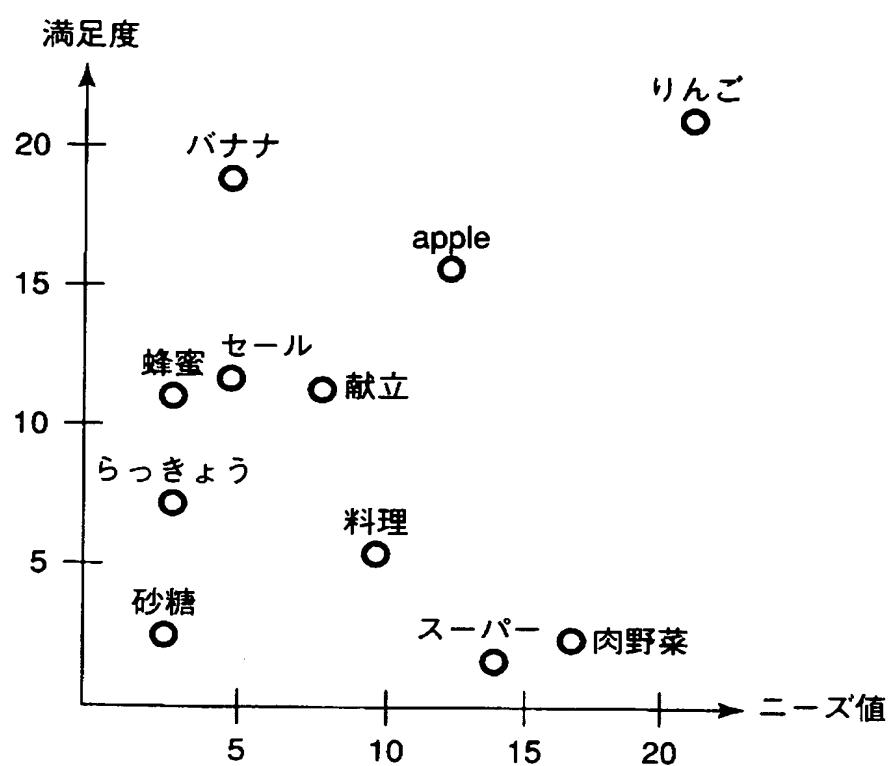
【図 6】



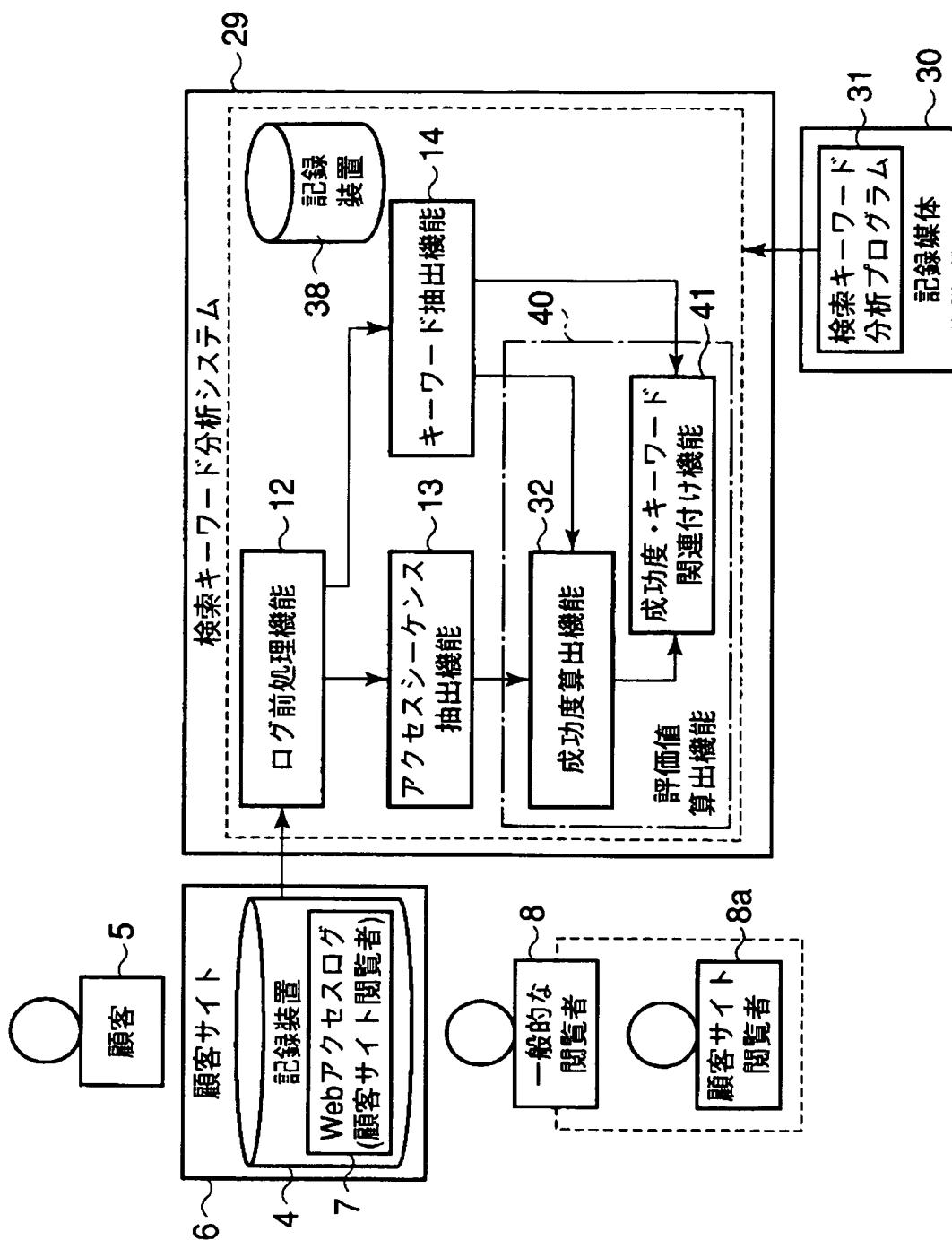
【図 7】



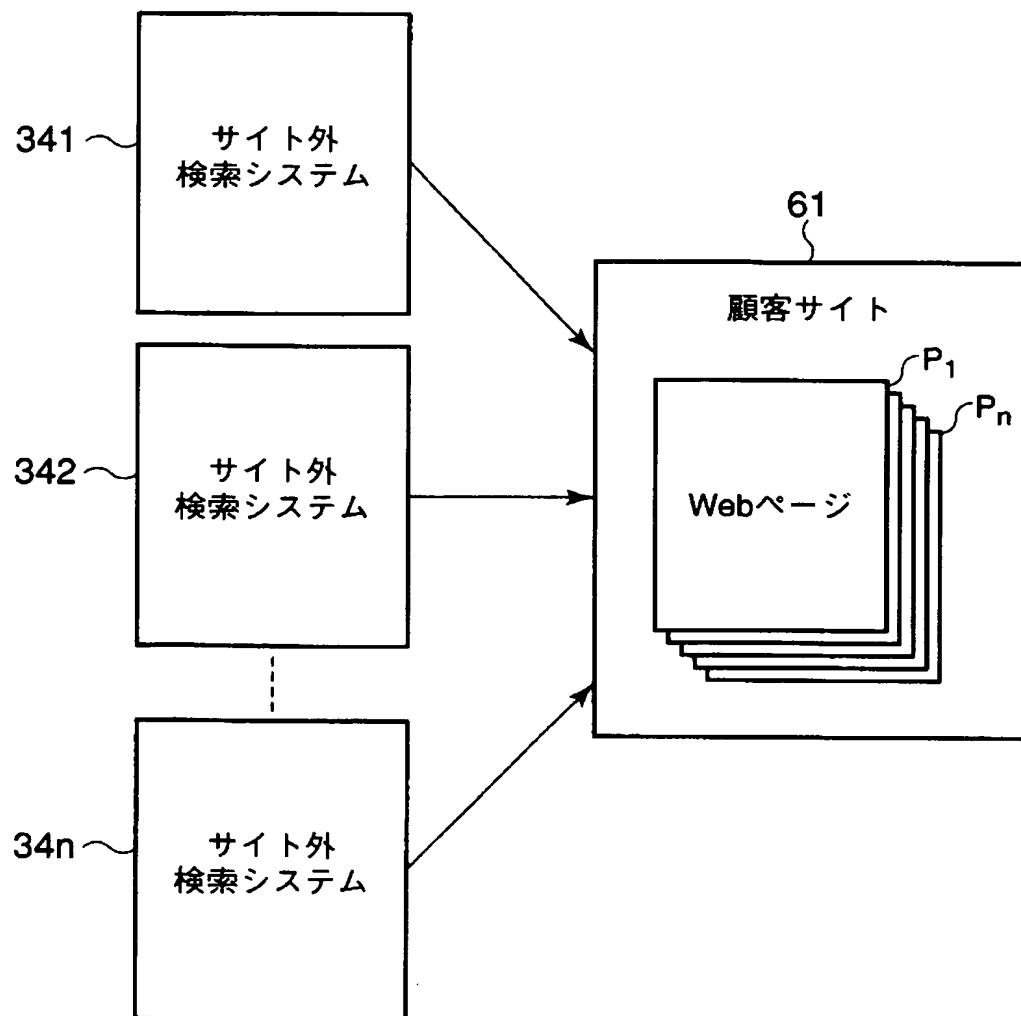
【図 8】



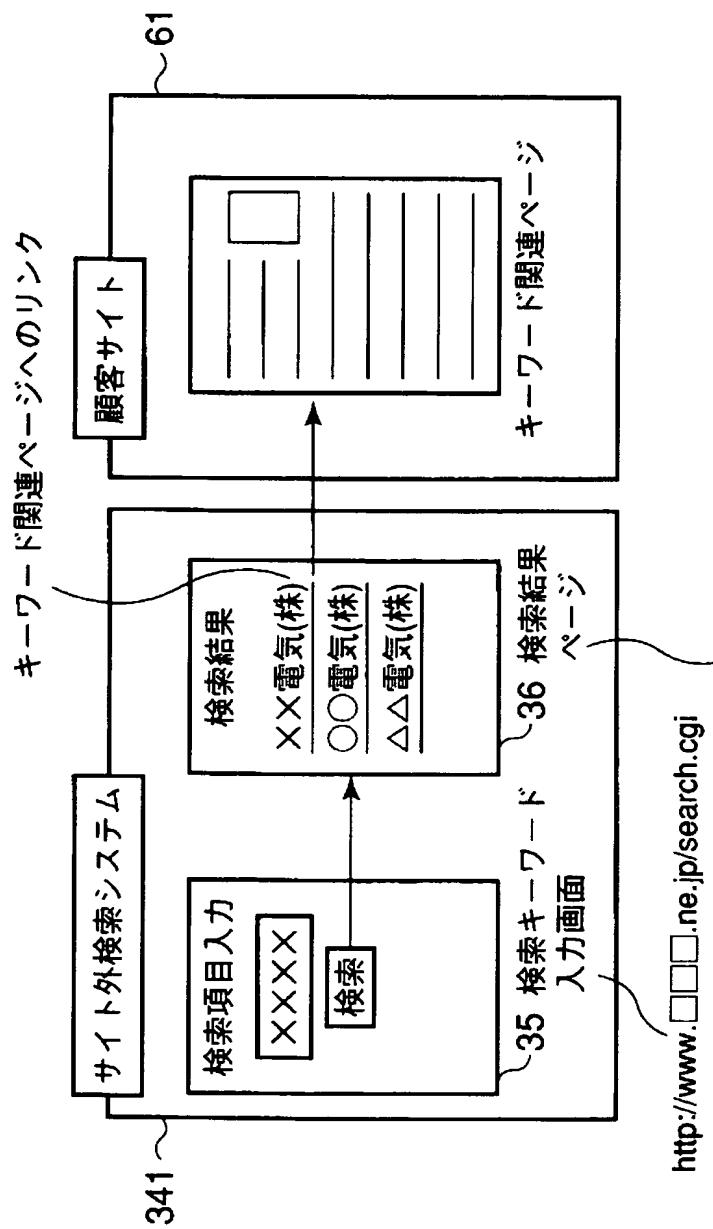
【図 9】



【図10】

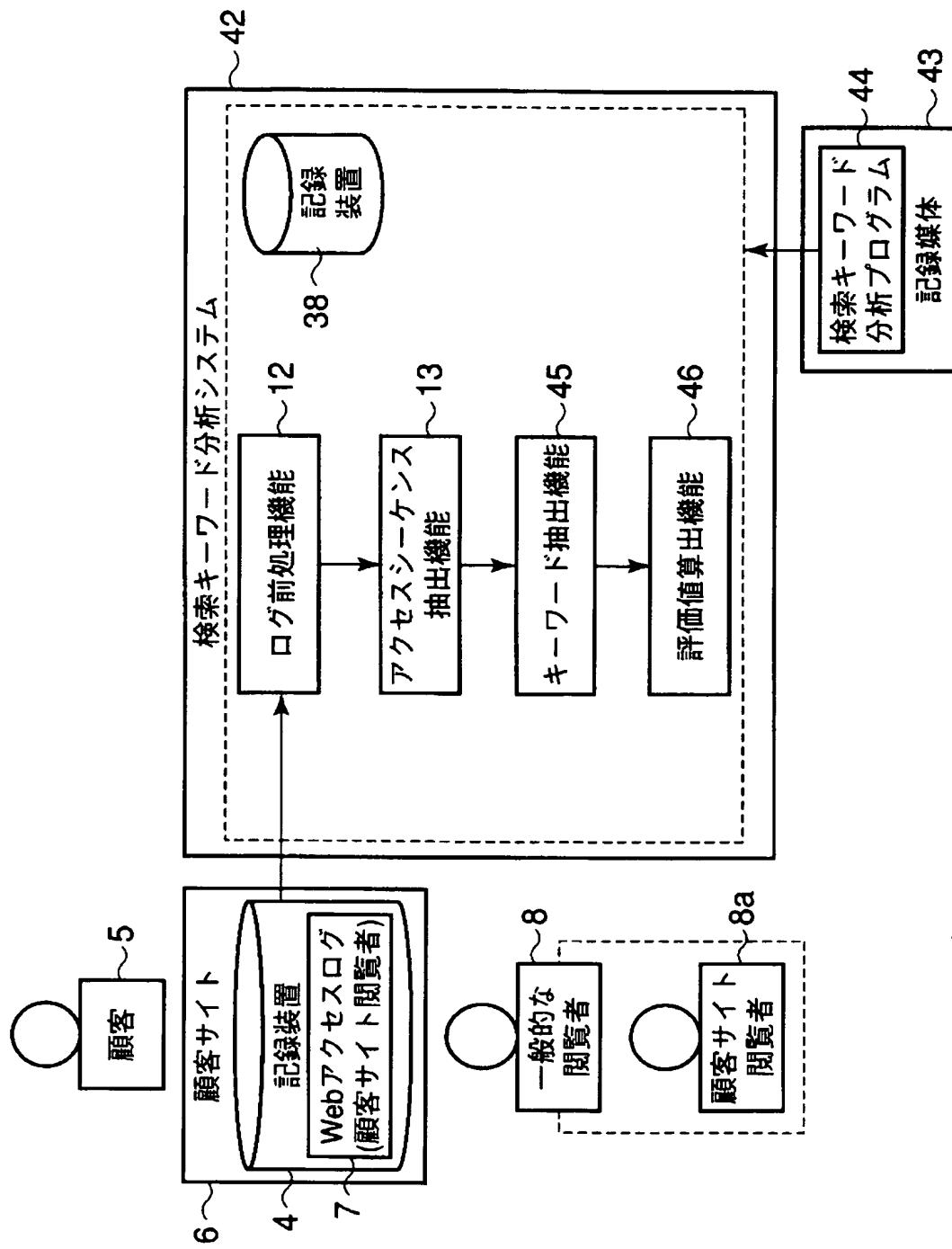


【図11】

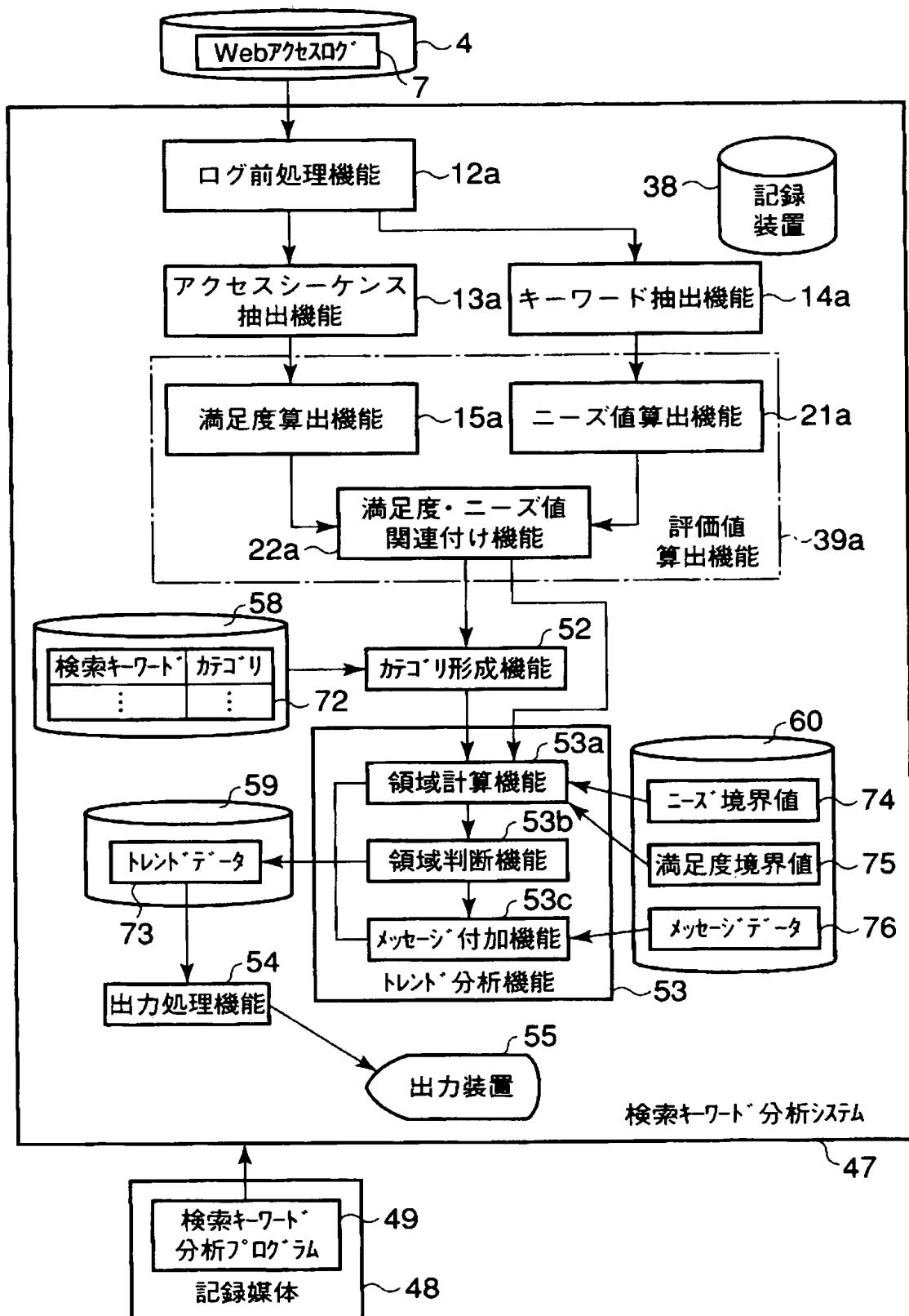


http://www.□□□.ne.jp/search.cgi?kw=XXXX(レファラ情報として顧客サイトのWebアクセスログに残る)

【図12】



【図13】



特願 2002-382396

出願人履歴情報

識別番号 [000003078]

1. 変更年月日 1990年 8月22日

[変更理由] 新規登録

住 所 神奈川県川崎市幸区堀川町72番地
氏 名 株式会社東芝

2. 変更年月日 2001年 7月 2日

[変更理由] 住所変更

住 所 東京都港区芝浦一丁目1番1号
氏 名 株式会社東芝

認定・付加情報（続巻）

【氏名又は名称】 河野 哲
【選任した代理人】
【識別番号】 100088683
【住所又は居所】 東京都千代田区霞が関3丁目7番2号 鈴榮特許
綜合法律事務所内
【氏名又は名称】 中村 誠
【選任した代理人】
【識別番号】 100070437
【住所又は居所】 東京都千代田区霞が関3丁目7番2号 鈴榮特許
綜合法律事務所内
【氏名又は名称】 河井 将次

次頁無

認定・付加情報

特許出願の番号	特願 2002-382396
受付番号	50201992828
書類名	特許願
担当官	第七担当上席 0096
作成日	平成 15 年 1 月 8 日

<認定情報・付加情報>

【特許出願人】

【識別番号】	000003078
【住所又は居所】	東京都港区芝浦一丁目 1 番 1 号
【氏名又は名称】	株式会社東芝

【代理人】

【識別番号】	100058479
【住所又は居所】	東京都千代田区霞が関 3 丁目 7 番 2 号 鈴榮特許 綜合法律事務所内
【氏名又は名称】	鈴江 武彦

【選任した代理人】

【識別番号】	100084618
【住所又は居所】	東京都千代田区霞が関 3 丁目 7 番 2 号 鈴榮特許 綜合法律事務所内
【氏名又は名称】	村松 貞男

【選任した代理人】

【識別番号】	100068814
【住所又は居所】	東京都千代田区霞が関 3 丁目 7 番 2 号 鈴榮特許 綜合法律事務所内
【氏名又は名称】	坪井 淳

【選任した代理人】

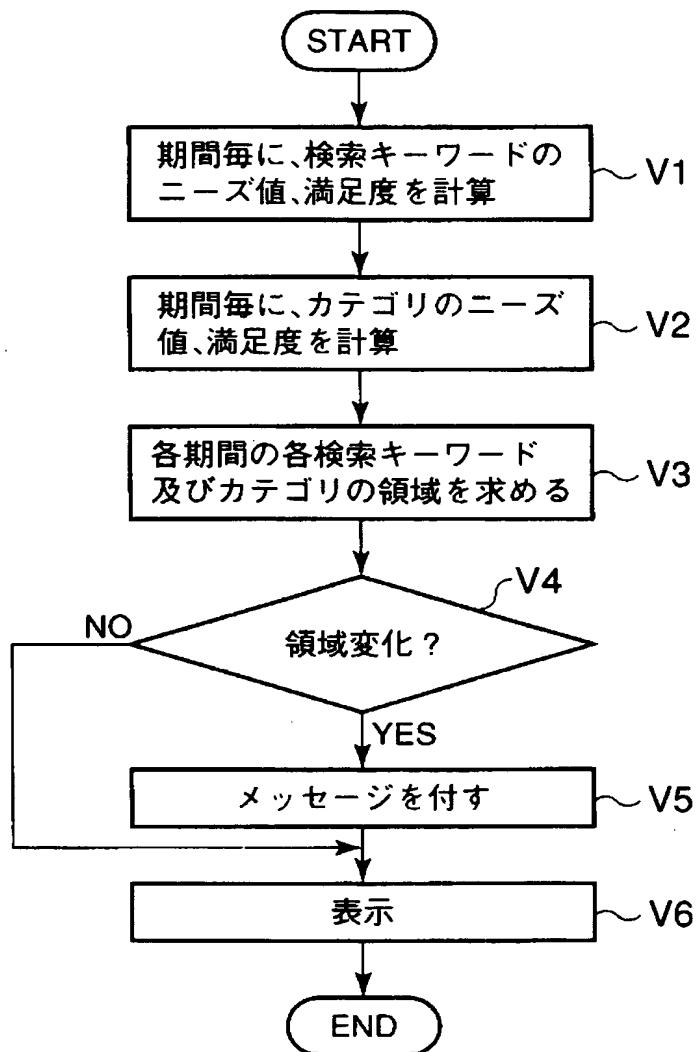
【識別番号】	100092196
【住所又は居所】	東京都千代田区霞が関 3 丁目 7 番 2 号 鈴榮特許 綜合法律事務所内
【氏名又は名称】	橋本 良郎

【選任した代理人】

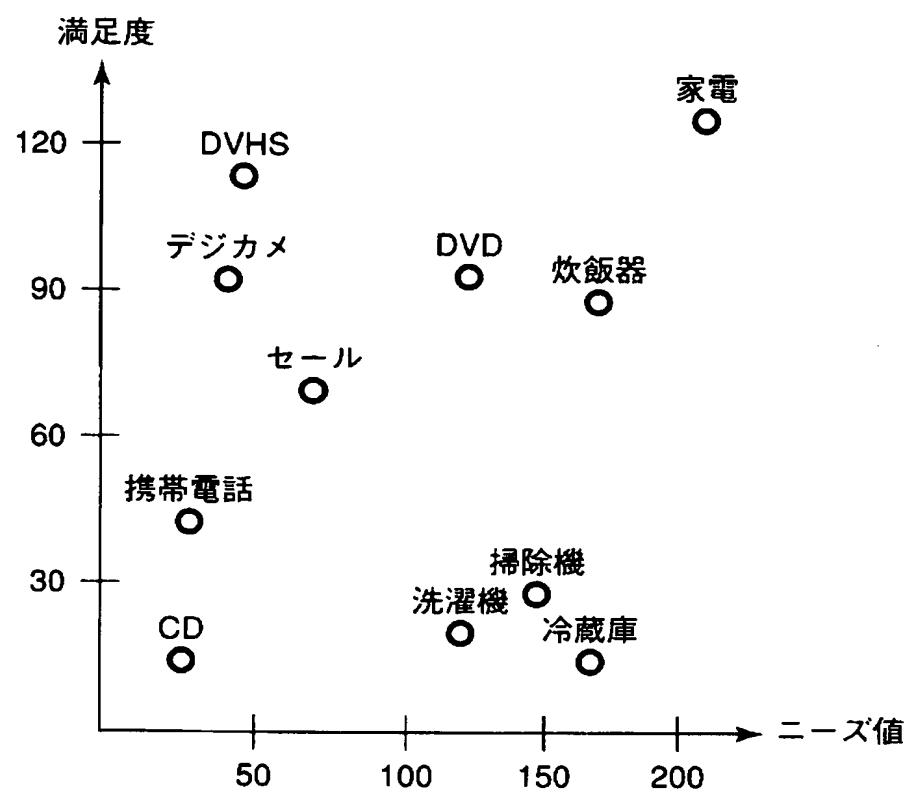
【識別番号】	100091351
【住所又は居所】	東京都千代田区霞が関 3 丁目 7 番 2 号 鈴榮特許 綜合法律事務所内

次頁有

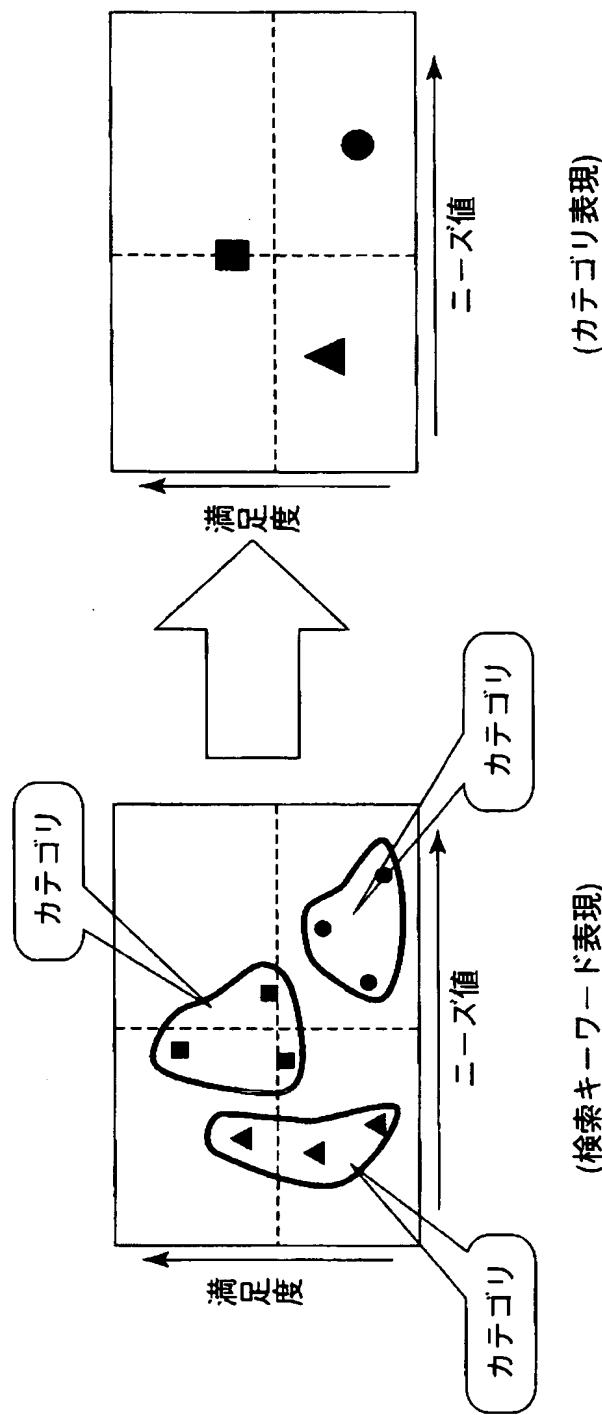
【図14】



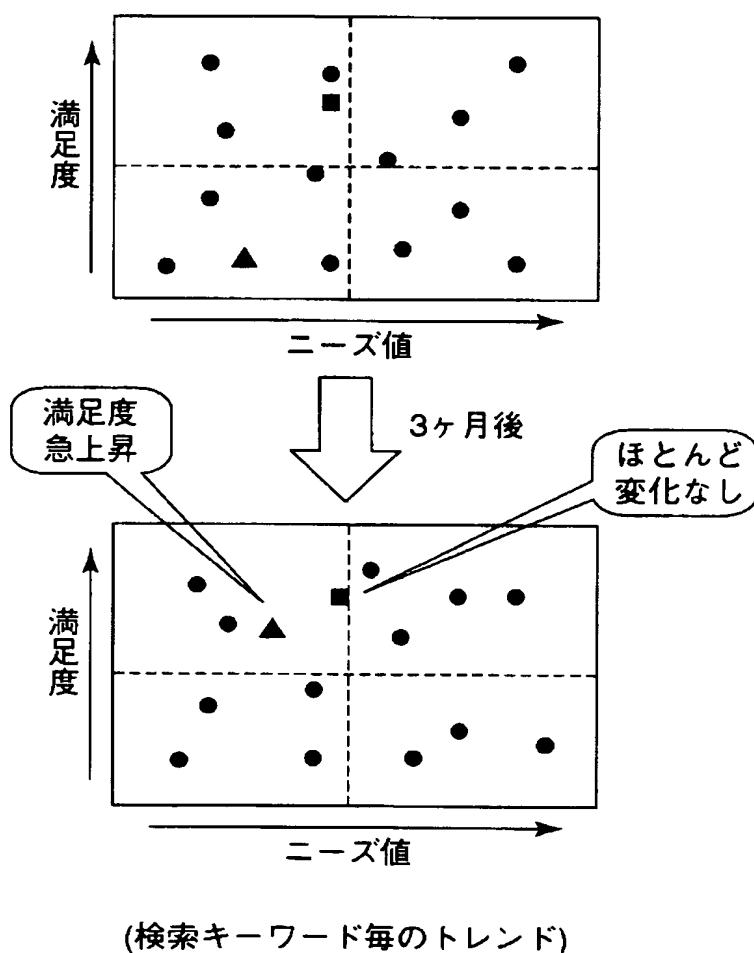
【図15】



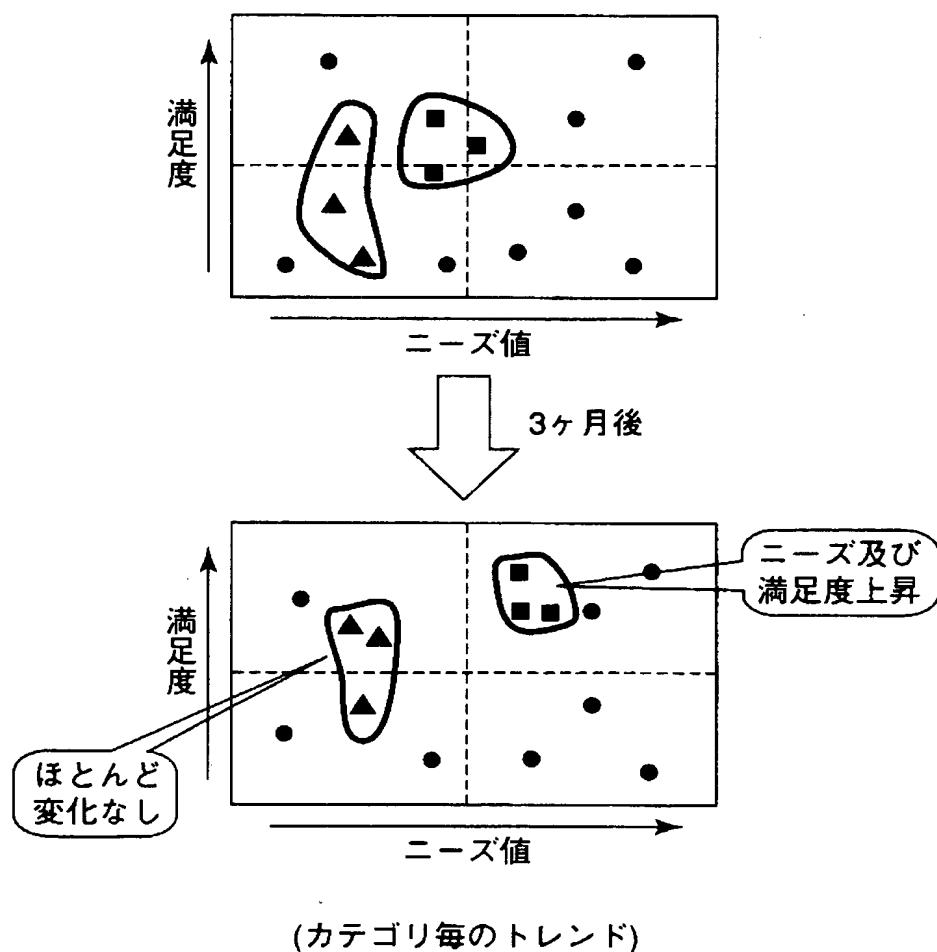
【図16】



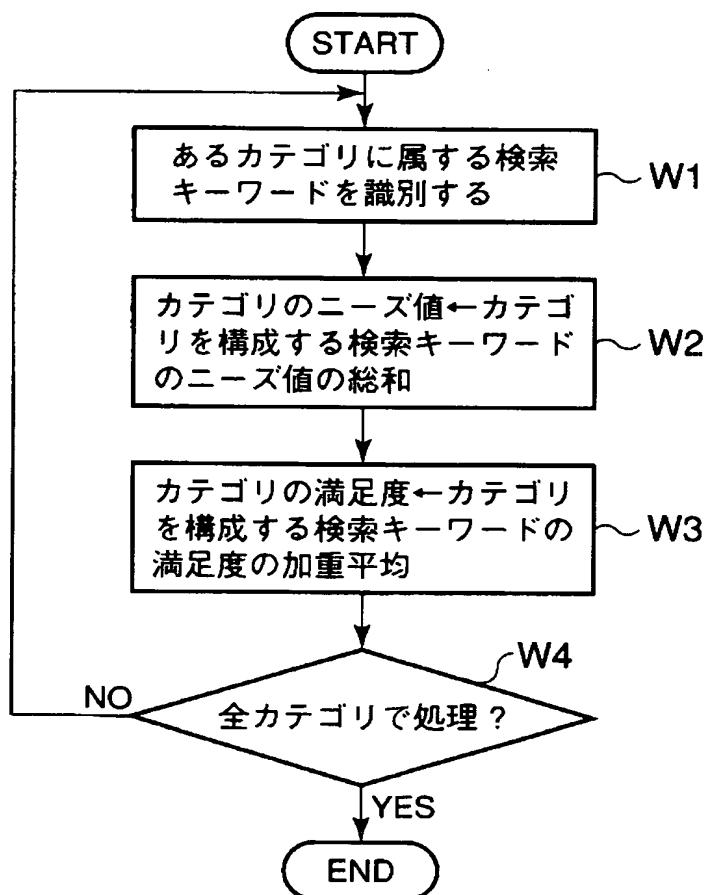
【図17】



【図18】



【図19】



【図20】

72 カテゴリデータ

キーワード	カテゴリ
食べ放題	グルメ
宴会	グルメ
ランチ	グルメ
せんそく	健康
ハウスダスト	健康

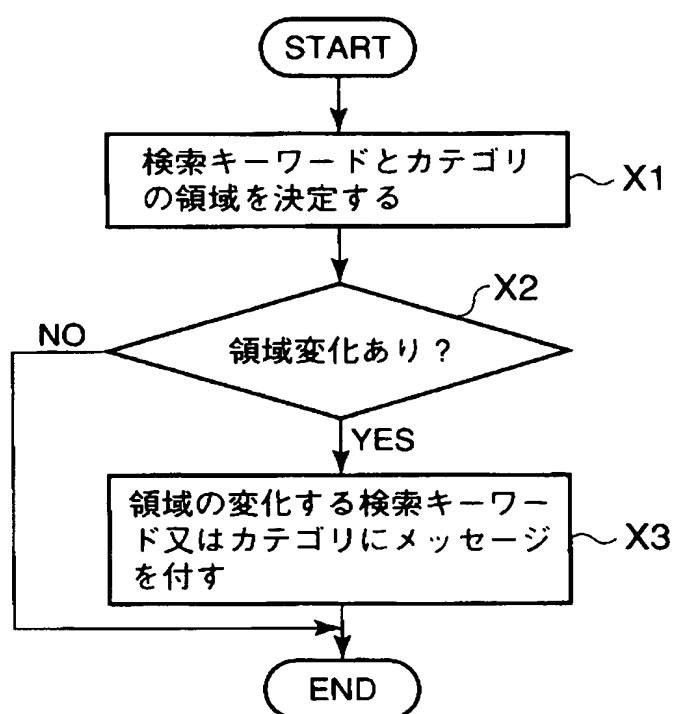
77 関係データ

キーワード	ニーズ	満足度
食べ放題	155	169
宴会	93	638
ランチ	124	221
せんそく	77	226
ハウスダスト	177	106

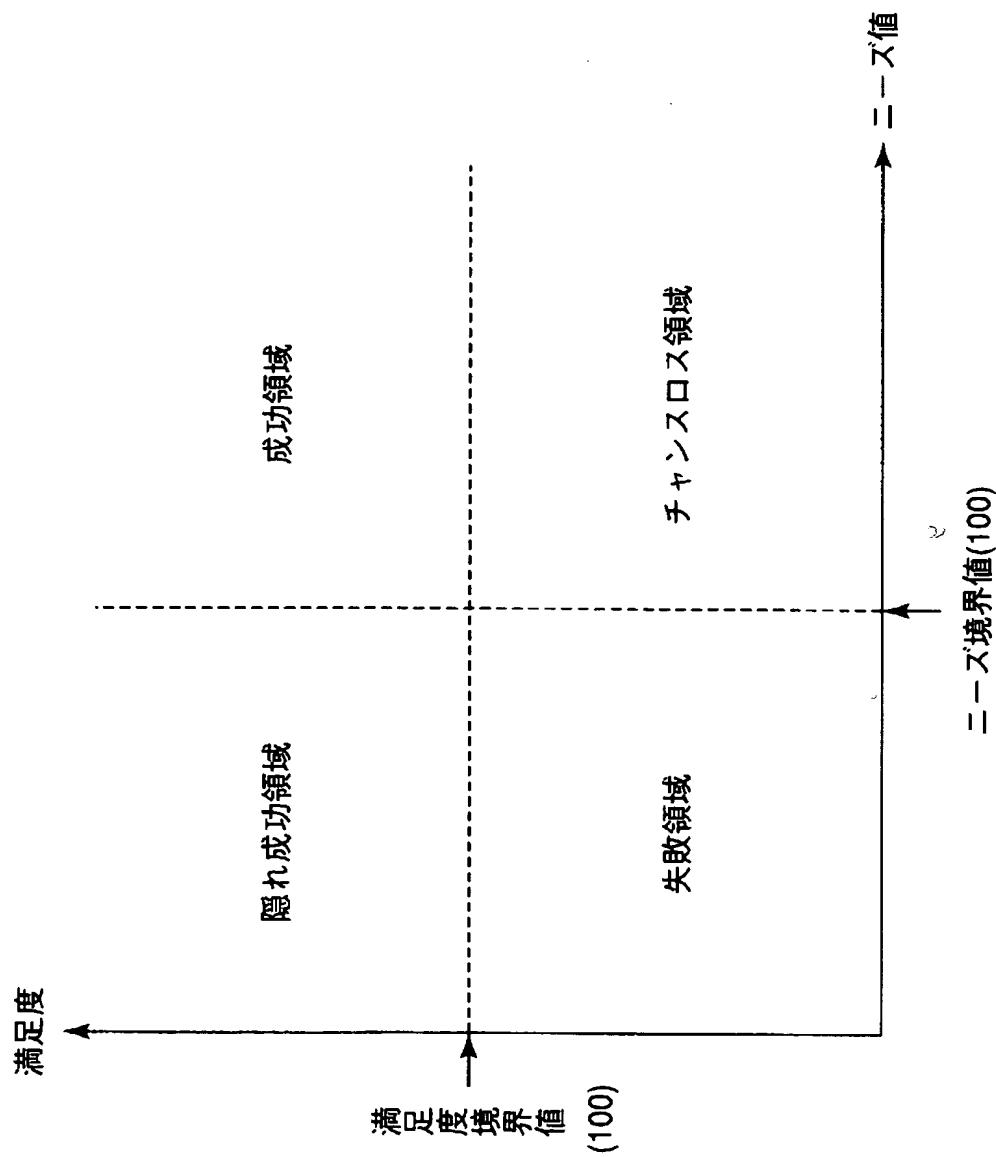
78 カテゴライズデータ

カテゴリ	ニーズ	満足度
グルメ	372	304
健康	254	142

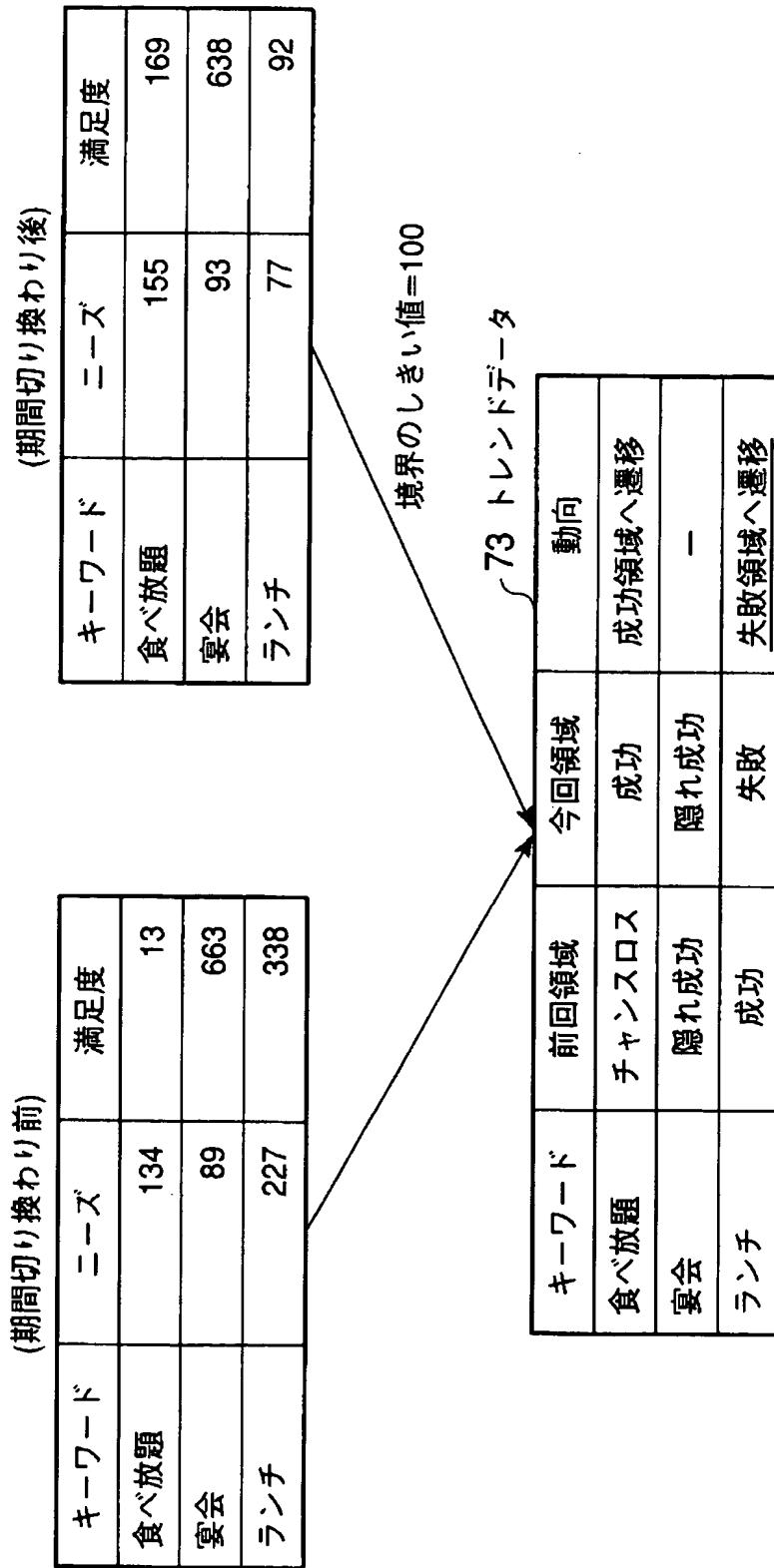
【図21】



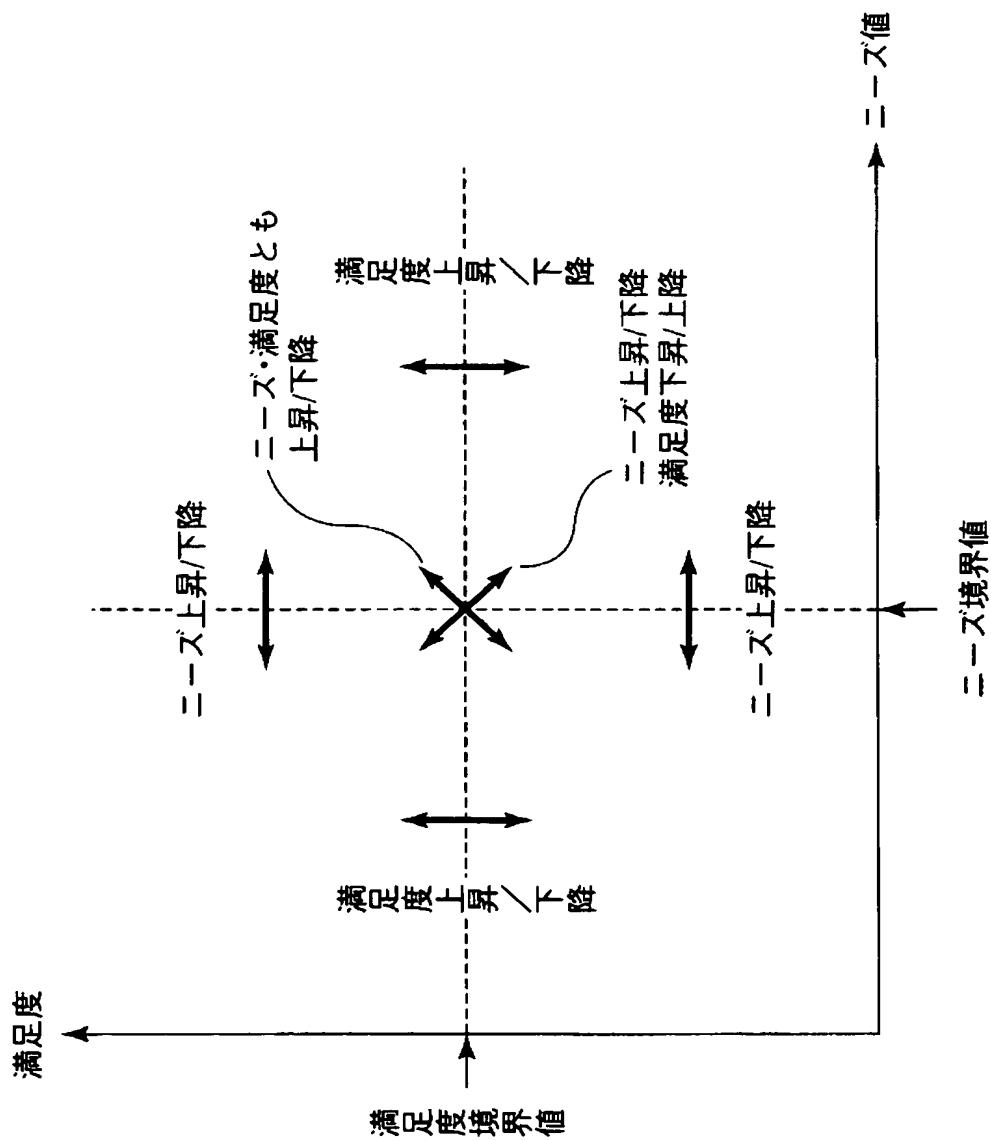
【図22】



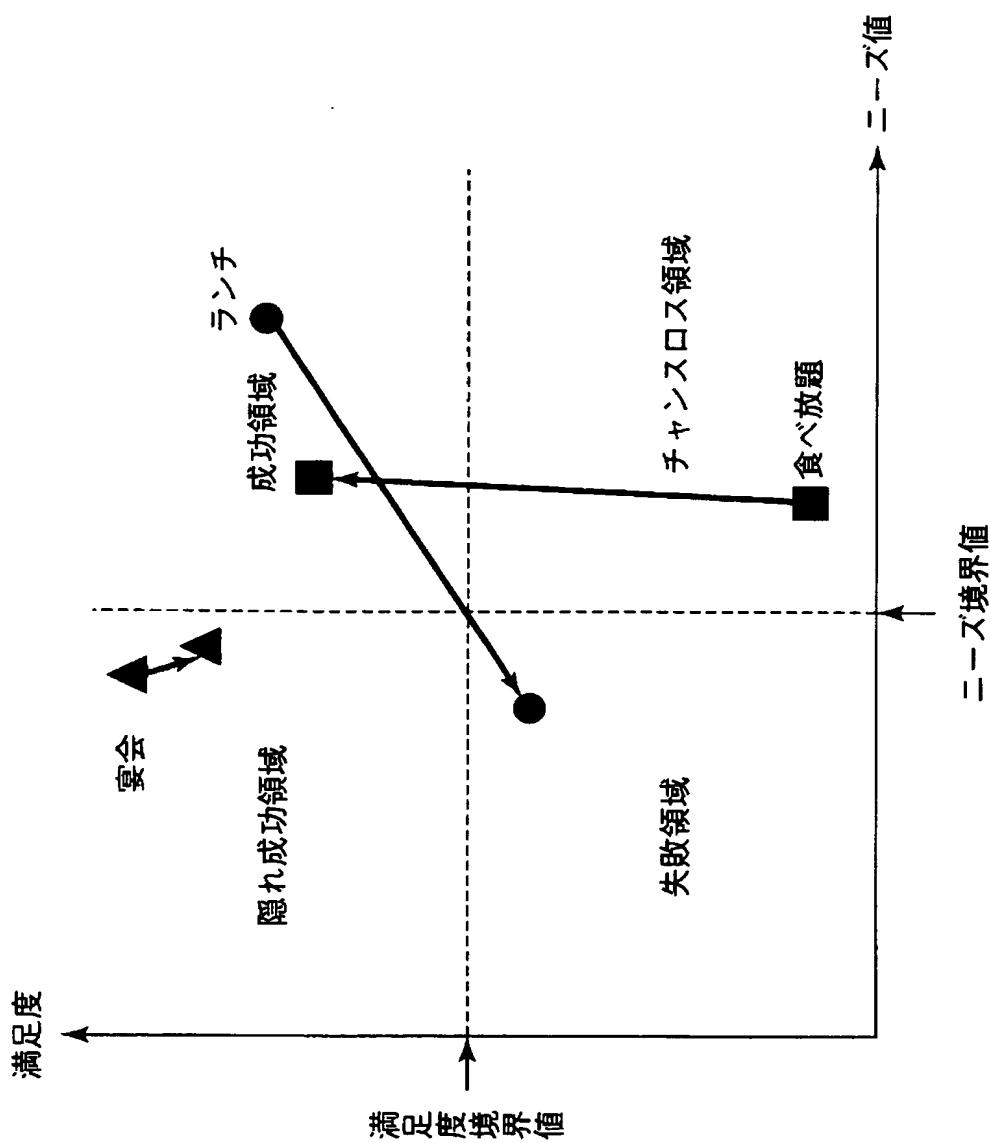
【図23】



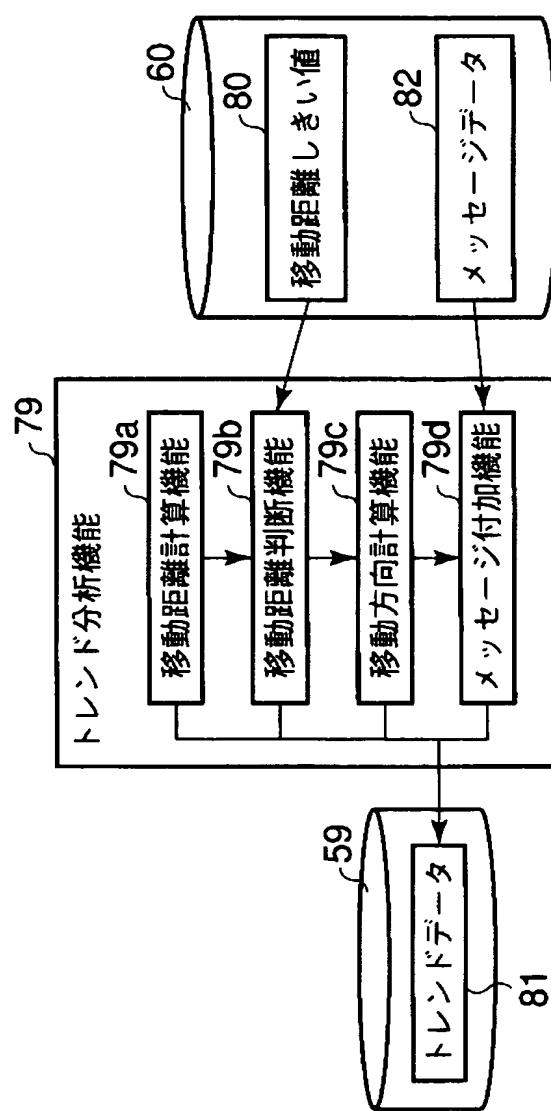
【図24】



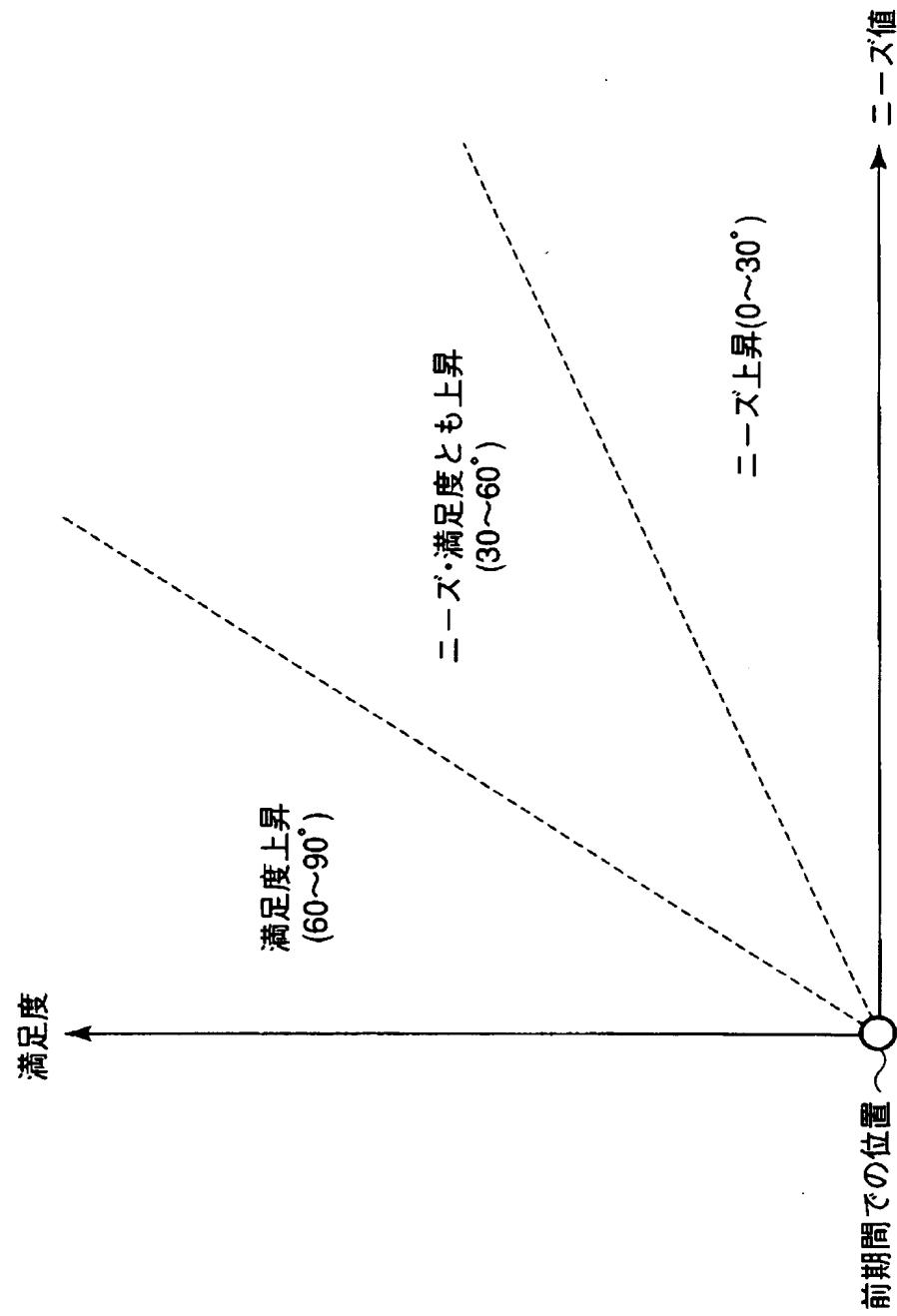
【図25】



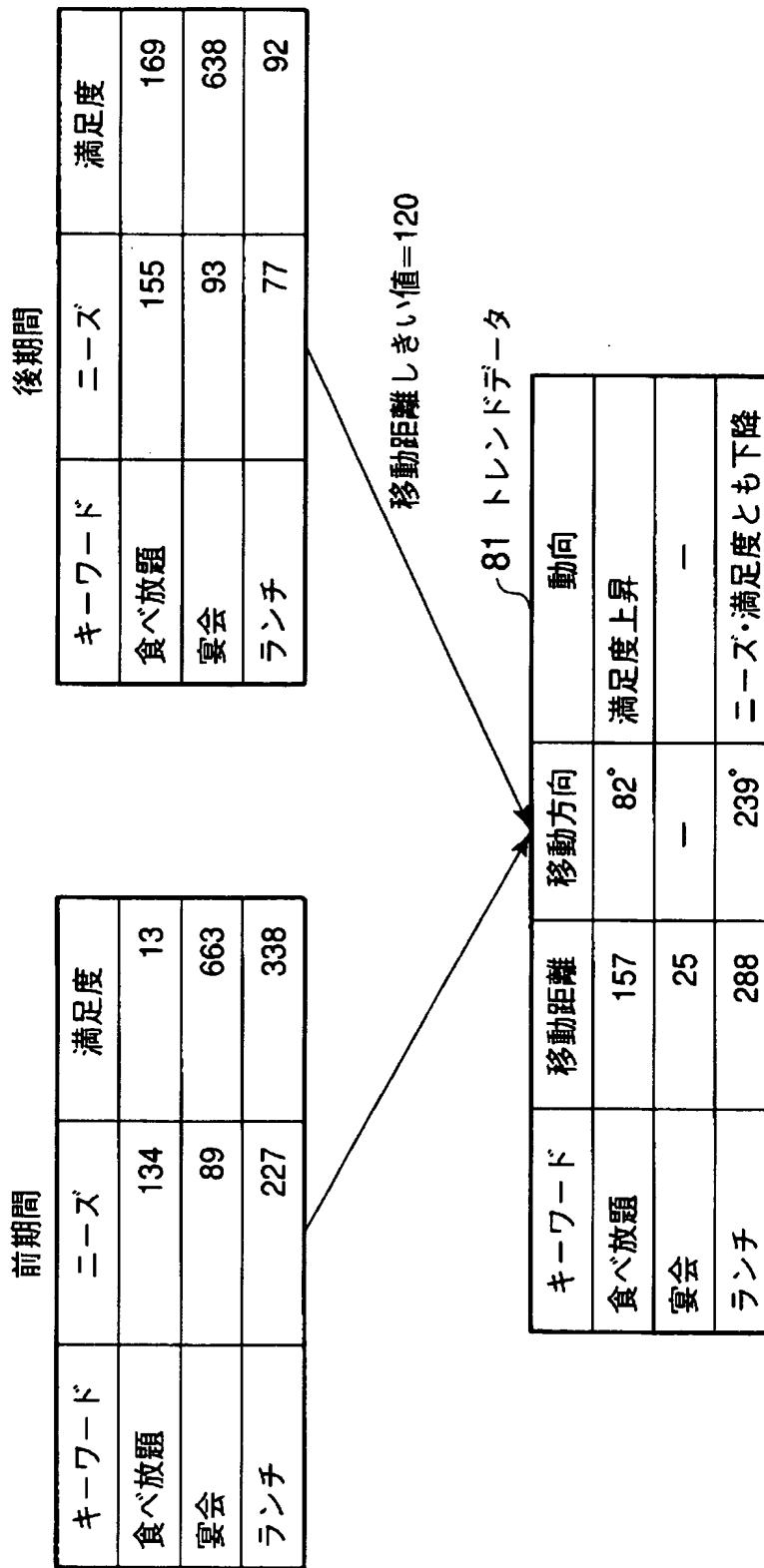
【図26】



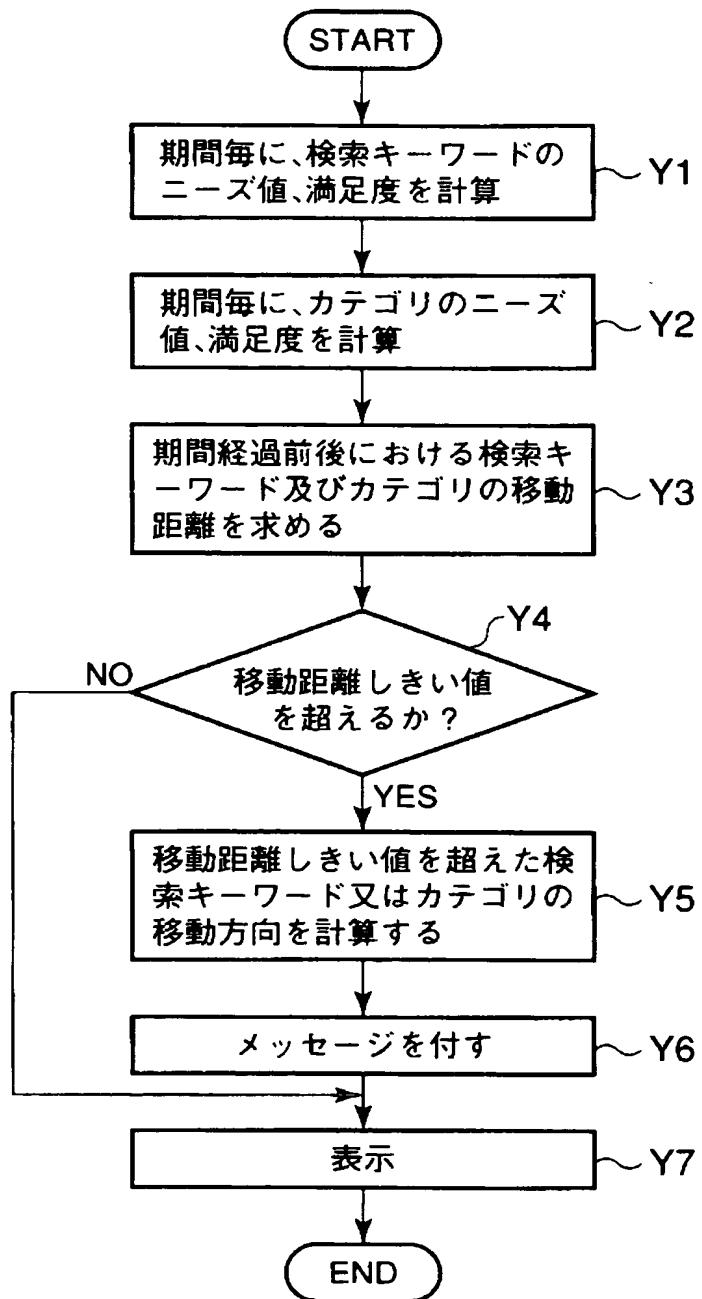
【図27】



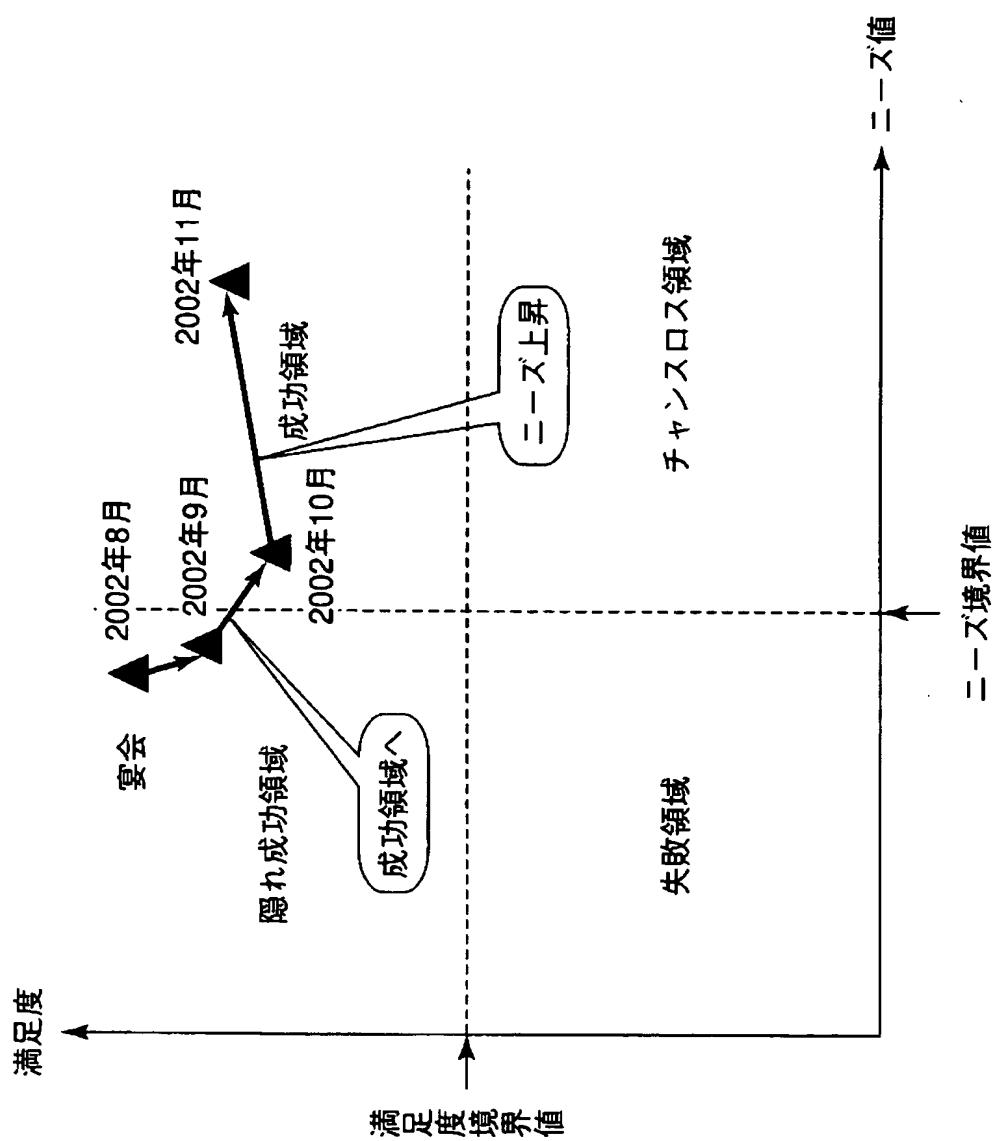
【図28】



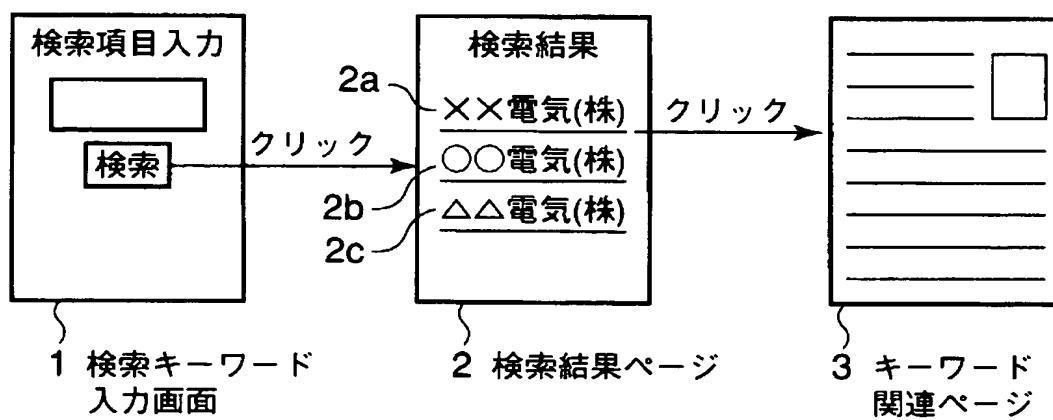
【図29】



【図30】



【図31】



【書類名】 要約書

【要約】

【課題】 入力された検索キーワードに基づく検索システムの検索結果に対する質的な評価値を求める。

【解決手段】 本発明のプログラム11は、コンピュータに、アクセス要求側を示すアクセス要求側識別データとアクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebログ7に基づいて同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出する機能13と、Webログ7において検索システムを示すアクセス先識別データと入力された検索キーワードとが関連付けされている場合に、検索キーワードを抽出する機能14と、アクセスの状態から検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいてアクセスシーケンスに対応する質的評価値を算出し、検索キーワードと質的評価値との対応関係を示す関係データを求める機能37とを実現させる。

【選択図】 図1

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 2003年 3月14日
Date of Application:

出願番号 特願2003-070281
Application Number:

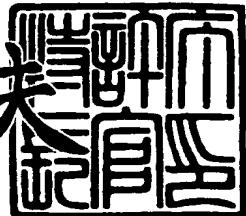
[ST. 10/C] : [JP 2003-070281]

出願人 株式会社東芝
Applicant(s):

2003年 9月 2日

特許庁長官
Commissioner,
Japan Patent Office

今井康夫



【書類名】 特許願
【整理番号】 A000300762
【提出日】 平成15年 3月14日
【あて先】 特許庁長官 殿
【国際特許分類】 G06F 17/00
【発明の名称】 検索キーワード分析プログラム及びシステム並びに方法
【請求項の数】 41
【発明者】
【住所又は居所】 東京都府中市東芝町1番地 株式会社東芝府中事業所内
【氏名】 平 博司
【発明者】
【住所又は居所】 東京都府中市東芝町1番地 株式会社東芝府中事業所内
【氏名】 飯田 晴彦
【発明者】
【住所又は居所】 東京都府中市東芝町1番地 株式会社東芝府中事業所内
【氏名】 平井 潤
【特許出願人】
【識別番号】 000003078
【氏名又は名称】 株式会社 東芝
【代理人】
【識別番号】 100058479
【弁理士】
【氏名又は名称】 鈴江 武彦
【電話番号】 03-3502-3181
【選任した代理人】
【識別番号】 100091351
【弁理士】
【氏名又は名称】 河野 哲

【選任した代理人】

【識別番号】 100088683

【弁理士】

【氏名又は名称】 中村 誠

【選任した代理人】

【識別番号】 100108855

【弁理士】

【氏名又は名称】 蔵田 昌俊

【選任した代理人】

【識別番号】 100084618

【弁理士】

【氏名又は名称】 村松 貞男

【選任した代理人】

【識別番号】 100092196

【弁理士】

【氏名又は名称】 橋本 良郎

【先の出願に基づく優先権主張】

【出願番号】 特願2002-243466

【出願日】 平成14年 8月23日

【先の出願に基づく優先権主張】

【出願番号】 特願2002-382396

【出願日】 平成14年12月27日

【手数料の表示】

【予納台帳番号】 011567

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9705037

【プルーフの要否】 要

・・・

【書類名】 明細書

【発明の名称】 検索キーワード分析プログラム及びシステム並びに方法

【特許請求の範囲】

【請求項1】 コンピュータに、

WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連續的なアクセスを表すアクセスシーケンスを抽出し、前記アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録するアクセスシーケンス抽出機能と、

前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けされている場合に、前記検索キーワードを抽出し、前記検索キーワードの抽出結果を示すキーワードデータを前記記録手段に記録するキーワード抽出機能と、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する質的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードのうち前記アクセスシーケンスにおいて入力された検索キーワードと前記質的評価値との対応関係を示す関係データを前記記録手段に記録する評価値算出機能と

を実現させるための検索キーワード分析プログラム。

【請求項2】 請求項1記載の検索キーワード分析プログラムにおいて、

前記評価値算出機能は、前記質的評価値に加えて、前記検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する量的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードのうち前記アクセスシーケンスにおいて入力された検索キ

ワードと前記質的評価値と前記量的評価値の対応関係を示す関係データを前記記録手段に記録する

ことを特徴とする検索キーワード分析プログラム。

【請求項 3】 コンピュータに、

WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、前記アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録するアクセスシーケンス抽出機能と、

前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けされており前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスにおいて前記検索キーワードが入力されている場合に、前記検索キーワードを抽出し、前記検索キーワードの抽出結果を示すキーワードデータを前記記録手段に記録するキーワード抽出機能と、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する質的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードと前記質的評価値との対応関係を示す関係データを前記記録手段に記録する評価値算出機能と

を実現させるための検索キーワード分析プログラム。

【請求項 4】 請求項 3 記載の検索キーワード分析プログラムにおいて、

前記評価値算出機能は、前記質的評価値に加えて、前記検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する量的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードと前記質的評価値と前記量的評価値の対応関係を示す関係データ

ータを前記記録手段に記録する
ことを特徴とする検索キーワード分析プログラム。

【請求項5】 請求項1又は請求項3記載の検索キーワード分析プログラムにおいて、

前記Webアクセスログは、レファラ情報と前記アクセス先識別データとを関連付けたデータを含み、

前記アクセスシーケンス抽出機能は、前記レファラ情報が前記検索システムを示す場合に、前記レファラ情報に関連付けされている前記アクセス先識別データの示すアクセス先へのアクセスを含むアクセスシーケンスを抽出し、

前記キーワード抽出機能は、前記レファラ情報が前記検索システムを示す場合に、前記レファラ情報に含まれている検索キーワードを抽出し、

前記評価値算出機能は、前記アクセスシーケンスに対する前記検索キーワード入力後の質的評価値を算出する

ことを特徴とする検索キーワード分析プログラム。

【請求項6】 コンピュータに、

WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとレファラ情報とを関連付けたWebアクセスログに基づいて、前記レファラ情報がWWWの検索システムを示す場合に前記レファラ情報に関連付けされている前記アクセス先識別データの示すアクセス先へのアクセスを含み同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、前記アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録するアクセスシーケンス抽出機能と、

前記レファラ情報が前記検索システムを示す場合に、前記レファラ情報に含まれている検索キーワードを抽出し、前記検索キーワードの抽出結果を示すキーワードデータを前記記録手段に記録するキーワード抽出機能と、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対する前記検索キ

一ワード入力後の質的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードのうち前記アクセスシーケンスに対応する検索キーワードと前記アクセスシーケンスに対応する質的評価値との対応関係を示す関係データを前記記録手段に記録する評価値算出機能とを実現させるための検索キーワード分析プログラム。

【請求項7】 請求項5又は請求項6記載の検索キーワード分析プログラムにおいて、

前記評価値算出機能は、前記質的評価値に加えて、前記検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対する前記検索キーワード入力後の量的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードと前記質的評価値と前記量的評価値の対応関係を示す関係データを前記記録手段に記録することを特徴とする検索キーワード分析プログラム。

【請求項8】 請求項2と請求項4と請求項7記載のいずれか1項記載の検索キーワード分析プログラムにおいて、

コンピュータに、前記記録手段に記録された関係データにしたがって、前記質的評価値と前記量的評価値とに基づくグラフを描画するグラフ化機能を実現させるためのプログラムを附加したことを特徴とする検索キーワード分析プログラム。

【請求項9】 請求項2と請求項4と請求項7と請求項8のいずれか1項記載の検索キーワード分析プログラムにおいて、

前記量的評価値は、前記検索キーワードの入力数が多いほど値が大きくなる検索結果のニーズ値であり、

前記量的評価規則は、前記検索キーワードの入力数が多いほど前記ニーズ値が増加する規則であることを特徴とする検索キーワード分析プログラム。

【請求項10】 請求項1乃至請求項9のいずれか1項記載の検索キーワード分析プログラムにおいて、

前記質的評価規則は、分析候補の検索キーワードと前記分析候補の検索キーワ

ードの入力後にアクセスされた場合に成功度を増加するアクセス先識別データとを関連付けたデータと前記記録手段に記録されたキーワードデータとに基づいて、前記アクセスシーケンスにおいて前記分析候補の検索キーワードが入力されておりその入力後前記分析候補の検索キーワードと関連付けされているアクセス先識別データの示すアクセス先がアクセスされている場合に、前記アクセスシーケンスに対する成功度を増加する規則であり、

前記質的評価値は、前記成功度であることを特徴とする検索キーワード分析プログラム。

【請求項11】 請求項10記載の検索キーワード分析プログラムにおいて

前記質的評価規則は、前記分析候補の検索キーワードが入力されてから前記分析候補の検索キーワードと関連付けされているアクセス先識別データの示すアクセス先がアクセスされるまでのアクセス数が少ないほど加えられる成功度を大きくすることを特徴とする検索キーワード分析プログラム。

【請求項12】 請求項1乃至請求項9のいずれか1項記載の検索キーワード分析プログラムにおいて、

前記Webアクセスログは、WWWにおけるアクセス毎に、前記アクセス要求側識別データと前記アクセス先識別データとアクセス時刻を示すデータとを関係付けており、

前記質的評価規則は、前記アクセスシーケンスにおける前記検索キーワード入力後のアクセスから最後のアクセスまでの時間間隔が長いほど前記アクセス要求側の満足度が増加する規則であり、

前記質的評価値は、前記満足度であることを特徴とする検索キーワード分析プログラム。

【請求項13】 コンピュータに、

WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスのうち検索結果ページへのアクセスから定まる始端と、前記始端と同一のア

セス要求側による連続的なアクセスのうち前記始端以降のアクセスから定まる終端とに基づくサブアクセスシーケンスを抽出し、前記サブアクセスシーケンスの抽出結果を示すサブアクセスシーケンスデータを記録手段に記録するサブアクセスシーケンス抽出機能と、

前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けされている場合に、前記検索キーワードを抽出し、前記検索キーワードの抽出結果を示すキーワードデータを前記記録手段に記録するキーワード抽出機能と、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記記録手段に記録されたサブアクセスシーケンスデータの示すサブアクセスシーケンスに対応する質的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードのうち前記サブアクセスシーケンスに対応する検索キーワードと前記質的評価値との対応関係を示す関係データを前記記録手段に記録する評価値算出機能と

を実現させるための検索キーワード分析プログラム。

【請求項14】 コンピュータに、

WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスから、検索結果ページへのアクセスから定まる始端と次の始端とを求めるとともに前記次の始端の直前のアクセスに基づいて前記始端に対応する終端を求めてサブアクセスシーケンスを抽出し、前記サブアクセスシーケンスの抽出結果を示すサブアクセスシーケンスデータを記録手段に記録するサブアクセスシーケンス抽出機能と、

前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けされている場合に、前記検索キーワードを抽出し、前記検索キーワード

ードの抽出結果を示すキーワードデータを前記記録手段に記録するキーワード抽出機能と、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記記録手段に記録されたサブアクセスシーケンスデータの示すサブアクセスシーケンスに対応する質的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードのうち前記サブアクセスシーケンスに対応する検索キーワードと前記質的評価値との対応関係を示す関係データを前記記録手段に記録する評価値算出機能と

を実現させるための検索キーワード分析プログラム。

【請求項15】 コンピュータに、

WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスから検索結果ページへのアクセスから定まる始端と次の始端とが求められた場合、前記次の始端の直前のアクセスを前記始端に対応する終端とし、前記始端が求められ前記次の始端が求められなかった場合、前記連続的なアクセスの終端を前記始端に対応する終端とし、前記始端と前記終端とに基づいてサブアクセスシーケンスを抽出し、前記サブアクセスシーケンスの抽出結果を示すサブアクセスシーケンスデータを記録手段に記録するサブアクセスシーケンス抽出機能と、

前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けされている場合に、前記検索キーワードを抽出し、前記検索キーワードの抽出結果を示すキーワードデータを前記記録手段に記録するキーワード抽出機能と、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記記録手段に記録されたサブアクセスシーケンスデータの示すサブアクセスシーケンスに対応する質的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検

索キーワードのうち前記サブアクセスシーケンスに対応する検索キーワードと前記質的評価値との対応関係を示す関係データを前記記録手段に記録する評価値算出機能と

を実現させるための検索キーワード分析プログラム。

【請求項16】 請求項2と請求項4と請求項7のいずれか1項記載の検索キーワード分析プログラムにおいて、

前記Webアクセスログは、前記アクセス要求側識別データと前記アクセス先識別データとアクセスされた時間を示す時間データとを関係付けたデータであり

前記評価値算出機能は、前記時間データに基づいて、複数の期間毎の前記関係データを求め、

コンピュータに、

前記複数の期間毎の関係データに基づいて、前記検索キーワードについて前記質的評価値と前記量的評価値との遷移状況を表すトレンドデータを求め、前記記録手段に記録するトレンド分析機能

を実現させることを特徴とする検索キーワード分析プログラム。

【請求項17】 請求項16記載の検索キーワード分析プログラムにおいて

前記トレンド分析機能は、前記複数の期間の切り換わりに応じて、前記検索キーワードの前記質的評価値と前記量的評価値とのうち少なくとも一方が、前記質的評価値と前記量的評価値のそれぞれについて設定された境界値を超えて変化するか判断し、

前記トレンドデータは、前記トレンド分析機能による判断結果を示すデータを含む

ことを特徴とする検索キーワード分析プログラム。

【請求項18】 請求項17記載の検索キーワード分析プログラムにおいて

前記トレンド分析機能は、

前記複数の期間毎に、前記複数の境界値から定まる複数の領域のうち、前記検

索キーワードの属する領域を求める領域計算機能と、

前記複数の期間の切り換わりに応じて前記検索キーワードの領域が変化するか判断する領域判断機能と、

前記領域が変化する検索キーワードに対して前記領域の変化に応じたメッセージを付す領域メッセージ付加機能と

を具備し、

前記トレンドデータは、前記領域が変化する検索キーワードに前記メッセージを付した結果を示すデータを含む

ことを特徴とする検索キーワード分析プログラム。

【請求項19】 請求項16乃至請求項18のいずれか1項記載の検索キーワード分析プログラムにおいて、

前記トレンド分析機能は、前記複数の期間毎の前記検索キーワードの前記質的評価値と前記量的評価値に基づいて、前記複数の期間の切り換わり前後における前記検索キーワードの移動距離と移動方向とを求め、

前記トレンドデータは、前記複数の期間の切り換わり前後における前記検索キーワードの移動距離と移動方向を示すデータを含むことを特徴とする検索キーワード分析プログラム。

【請求項20】 請求項16乃至請求項19のいずれか1項記載の検索キーワード分析プログラムにおいて、

前記トレンド分析機能は、

前記複数の期間毎の前記検索キーワードの前記質的評価値と前記量的評価値に基づいて、前記複数の期間の切り換わり前後における前記検索キーワードの移動距離を求める移動距離計算機能と、

前記移動距離が移動距離しきい値を超えているか判断する移動距離判断機能と

前記移動距離が前記移動距離しきい値を超えている場合に、前記検索キーワードの移動方向を求める移動方向計算機能と、

前記移動距離が前記移動距離しきい値を超える検索キーワードに対して前記移動方向に応じたメッセージを付すベクトルメッセージ付加機能と

を具備し、

前記トレンドデータは、前記移動距離が前記移動距離しきい値を超える検索キーワードに前記メッセージを付した結果を示すデータを含むことを特徴とする検索キーワード分析プログラム。

【請求項21】 請求項16乃至請求項20のいずれか1項記載の検索キーワード分析プログラムにおいて、

コンピュータに、

前記トレンドデータに基づいて、前記質的評価値と前記量的評価値とを複数の座標軸とする空間上に、前記複数の期間毎の前記検索キーワードを示す指標を前記質的評価値と前記量的評価値とに応じて配置し、配置された前記指標を前記複数の期間の経過順序に応じて線で結んだグラフを表示するための処理を実行する出力処理機能を実現させる

ことを特徴とする検索キーワード分析プログラム。

【請求項22】 請求項2と請求項4と請求項7のいずれか1項記載の検索キーワード分析プログラムにおいて、

前記Webアクセスログは、前記アクセス要求側識別データと前記アクセス先識別データとアクセスされた時間を示す時間データとを関係付けたデータであり、

前記評価値算出機能は、前記時間データに基づいて、複数の期間毎の前記関係データを求め、

コンピュータに、

前記検索キーワードと前記検索キーワードの属するカテゴリとを関係付けたカテゴリデータと前記複数の期間毎の関係データとに基づいて、前記複数の期間毎に前記カテゴリについて前記質的評価値と前記量的評価値とを算出し、前記記録手段に記録するカテゴリ形成機能と、

前記カテゴリについて前記質的評価値と前記量的評価値との遷移状況を表すトレンドデータを求め、前記記録手段に記録するトレンド分析機能とを実現させることを特徴とする検索キーワード分析プログラム。

【請求項23】 請求項22記載の検索キーワード分析プログラムにおいて

コンピュータに、

前記トレンドデータに基づいて、前記質的評価値と前記量的評価値とを複数の座標軸とする空間上に、前記複数の期間毎の前記カテゴリを示す指標を前記質的評価値と前記量的評価値とに応じて配置し、配置された前記指標を前記複数の期間の経過順序に応じて線で結んだグラフを表示するための処理を実行する出力処理機能を実現させる

ことを特徴とする検索キーワード分析プログラム。

【請求項24】 WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出するアクセスシーケンス抽出手段と、

前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けされている場合に、前記検索キーワードを抽出するキーワード抽出手段と、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記アクセスシーケンスに対応する質的評価値を算出し、前記アクセスシーケンスにおいて入力された検索キーワードと前記質的評価値との対応関係を示す関係データを作成する評価値算出手段と

を具備する検索キーワード分析システム。

【請求項25】 請求項24記載の検索キーワード分析システムにおいて、前記評価値算出手段は、前記質的評価値に加えて、前記検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、前記アクセスシーケンスに対応する量的評価値を算出し、前記アクセスシーケンスにおいて入力された検索キーワードと前記質的評価値と前記量的評価値の対応関係を示す関係データを作成する

ことを特徴とする検索キーワード分析システム。

【請求項 26】 WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出するアクセスシーケンス抽出手段と、

前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けされており前記アクセスシーケンスにおいて前記検索キーワードが入力されている場合に、前記検索キーワードを抽出するキーワード抽出手段と、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記アクセスシーケンスに対応する質的評価値を算出し、前記検索キーワードと前記質的評価値との対応関係を示す関係データを作成する評価値算出手段と

を具備する検索キーワード分析システム。

【請求項 27】 請求項 26 記載の検索キーワード分析システムにおいて、前記評価値算出手段は、前記質的評価値に加えて、前記検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、前記アクセスシーケンスに対応する量的評価値を算出し、前記検索キーワードと前記質的評価値と前記量的評価値の対応関係を示す関係データを作成する

ことを特徴とする検索キーワード分析システム。

【請求項 28】 WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとレファラ情報とを関連付けたWebアクセスログに基づいて、前記レファラ情報がWWWの検索システムを示す場合に前記レファラ情報に関連付けされている前記アクセス先識別データの示すアクセス先へのアクセスを含み同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出するアクセスシーケンス抽出手段と、

前記レファラ情報が前記検索システムを示す場合に、前記レファラ情報に含ま

れている検索キーワードを抽出するキーワード抽出手段と、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記アクセスシーケンスに対する前記検索キーワード入力後の質的評価値を算出し、前記アクセスシーケンスに対応する前記検索キーワードと前記質的評価値との対応関係を示す関係データを作成する評価値算出手段と
を具備する検索キーワード分析システム。

【請求項29】 請求項28記載の検索キーワード分析システムにおいて、

前記評価値算出手段は、前記質的評価値に加えて、前記検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、前記アクセスシーケンスに対する前記検索キーワード入力後の量的評価値を算出し、前記検索キーワードと前記質的評価値と前記量的評価値の対応関係を示す関係データを作成する

ことを特徴とする検索キーワード分析システム。

【請求項30】 WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスのうち検索結果ページへのアクセスから定まる始端と、前記始端と同一のアクセス要求側による連続的なアクセスのうち前記始端以降のアクセスから定まる終端に基づくサブアクセスシーケンスを抽出するサブアクセスシーケンス抽出手段と、

前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けされている場合に、前記検索キーワードを抽出するキーワード抽出手段と、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記サブアクセスシーケンスに対応する質的評価値を算出し、前記アクセスシーケンスに対応する検索キーワードと前記質的評価値との対応関係を示す関係データを作成する評価値

算出手段と
を具備する検索キーワード分析システム。

【請求項31】 請求項25と請求項27と請求項29のいずれか1項記載の検索キーワード分析システムにおいて、

前記Webアクセスログは、前記アクセス要求側識別データと前記アクセス先識別データとアクセスされた時間を示す時間データとを関係付けたデータであり

前記評価値算出手段は、前記時間データに基づいて、複数の期間毎の前記関係データを求め、

前記複数の期間毎の関係データに基づいて、前記検索キーワードについて前記質的評価値と前記量的評価値との遷移状況を表すトレンドデータを求めるトレンド分析手段を具備することを特徴とする検索キーワード分析システム。

【請求項32】 請求項25と請求項27と請求項29のいずれか1項記載の検索キーワード分析システムにおいて、

前記Webアクセスログは、前記アクセス要求側識別データと前記アクセス先識別データとアクセスされた時間を示す時間データとを関係付けたデータであり

前記評価値算出手段は、前記時間データに基づいて、複数の期間毎の前記関係データを求め、

前記検索キーワードと前記検索キーワードの属するカテゴリとを関係付けたカテゴリデータと前記複数の期間毎の関係データとに基づいて、前記複数の期間毎に前記カテゴリについて前記質的評価値と前記量的評価値とを算出し、前記記録手段に記録するカテゴリ形成手段と、

前記カテゴリについて前記質的評価値と前記量的評価値との遷移状況を表すトレンドデータを求めるトレンド分析手段と
を具備する検索キーワード分析システム。

【請求項33】 WWWにおける検索システムに入力された検索キーワードについて分析するためにコンピュータにより実施される検索キーワード分析方法において、

WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、前記アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録するとともに、前記WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けされている場合に、前記検索キーワードを抽出し、前記検索キーワードの抽出結果を示すキーワードデータを前記記録手段に記録し、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する質的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードのうち前記アクセスシーケンスにおいて入力された検索キーワードと前記質的評価値との対応関係を示す関係データを前記記録手段に記録することを特徴とする検索キーワード分析方法。

【請求項34】 請求項33記載の検索キーワード分析方法において、前記質的評価値とともに、前記検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する量的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードのうち前記アクセスシーケンスにおいて入力された検索キーワードと前記質的評価値と前記量的評価値の対応関係を示す関係データを前記記録手段に記録することを特徴とする検索キーワード分析方法。

【請求項35】 WWWにおける検索システムに入力された検索キーワードについて分析するためにコンピュータにより実施される検索キーワード分析方法において、

WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付け

たWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、前記アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録し、

前記Webアクセスログにおいて前記検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けされており前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスにおいて前記検索キーワードが入力されている場合に、前記検索キーワードを抽出し、前記検索キーワードの抽出結果を示すキーワードデータを前記記録手段に記録し、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する質的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードと前記質的評価値との対応関係を示す関係データを前記記録手段に記録することを特徴とする検索キーワード分析方法。

【請求項36】 請求項35記載の検索キーワード分析方法において、

前記質的評価値に加えて、前記検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する量的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードと前記質的評価値と前記量的評価値の対応関係を示す関係データを前記記録手段に記録することを特徴とする検索キーワード分析方法。

【請求項37】 WWWにおける検索システムに入力された検索キーワードについて分析するためにコンピュータにより実施される検索キーワード分析方法において、

WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとレファラ情報とを関連付けたWebアクセスログに基づいて、前記レファラ情報が前記検索

システムを示す場合に前記レファラ情報に関連付けされている前記アクセス先識別データの示すアクセス先へのアクセスを含み同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、前記アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録するとともに、前記レファラ情報が前記検索システムを示す場合に、前記レファラ情報に含まれている検索キーワードを抽出し、前記検索キーワードの抽出結果を示すキーワードデータを前記記録手段に記録し、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対する前記検索キーワード入力後の質的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードのうち前記アクセスシーケンスに対応する検索キーワードと前記アクセスシーケンスに対応する質的評価値との対応関係を示す関係データを前記記録手段に記録する

ことを特徴とする検索キーワード分析方法。

【請求項38】 請求項37記載の検索キーワード分析方法において、

前記質的評価値とともに、前記検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、前記記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対する前記検索キーワード入力後の量的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードと前記質的評価値と前記量的評価値の対応関係を示す関係データを前記記録手段に記録する

ことを特徴とする検索キーワード分析方法。

【請求項39】 WWWにおける検索システムに入力された検索キーワードについて分析するためにコンピュータにより実施される検索キーワード分析方法において、

WWWにおけるアクセス要求側を示すアクセス要求側識別データと前記アクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセ

スのうち検索結果ページへのアクセスから定まる始端と、前記始端と同一のアクセス要求側による連続的なアクセスのうち前記始端以降のアクセスから定まる終端とに基づくサブアクセスシーケンスを抽出し、前記サブアクセスシーケンスの抽出結果を示すサブアクセスシーケンスデータを記録手段に記録するとともに、前記Webアクセスログにおいて前記検索システムを示すアクセス先識別データと前記アクセス要求側から前記検索システムに入力された検索キーワードとが関連付けされている場合に、前記検索キーワードを抽出し、前記検索キーワードの抽出結果を示すキーワードデータを前記記録手段に記録し、

WWWにおけるアクセスの状態から前記検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、前記記録手段に記録されたサブアクセスシーケンスデータの示すサブアクセスシーケンスに対応する質的評価値を算出し、前記記録手段に記録されているキーワードデータの示す検索キーワードのうち前記サブアクセスシーケンスに対応する検索キーワードと前記質的評価値との対応関係を示す関係データを前記記録手段に記録することを特徴とする検索キーワード分析方法。

【請求項40】 請求項34と請求項36と請求項38のいずれか1項記載の検索キーワード分析方法において、

前記Webアクセスログは、前記アクセス要求側識別データと前記アクセス先識別データとアクセスされた時間を示す時間データとを関係付けたデータであり

前記関係データは、前記時間データに基づいて、複数の期間毎に求められ、前記複数の期間毎の関係データに基づいて、前記検索キーワードについて前記質的評価値と前記量的評価値との遷移状況を表すトレンドデータを求め、前記記録手段に記録する

ことを特徴とする検索キーワード分析方法。

【請求項41】 請求項34と請求項36と請求項38のいずれか1項記載の検索キーワード分析方法において、

前記Webアクセスログは、前記アクセス要求側識別データと前記アクセス先識別データとアクセスされた時間を示す時間データとを関係付けたデータであり

前記関係データは、前記時間データに基づいて、複数の期間毎に求められ、前記検索キーワードと前記検索キーワードの属するカテゴリとを関係付けたカテゴリデータと前記複数の期間毎の関係データとにに基づいて、前記複数の期間毎に前記カテゴリについて前記質的評価値と前記量的評価値とを算出し、前記記録手段に記録し、

前記カテゴリについて前記質的評価値と前記量的評価値との遷移状況を表すトレンドデータを求め、前記記録手段に記録することを特徴とする検索キーワード分析方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、WWW (World Wide Web) において、目的のWebページ又はWebサイトを探すために、Web閲覧者（以下、単に「閲覧者」という）からWWW検索システム（以下、「検索システム」という）に入力された検索キーワードの特徴を分析するためのプログラム及びシステム並びに方法に関する。

【0002】

【従来の技術】

インターネットにおけるアクセス先の例として、Webページ又はWebサイトがある。

【0003】

インターネット上のWebページ又はWebサイトは、宣伝広告、情報発信、受注受付などの様々な目的のために運営される。かかる目的を効果的に達成するために、Webページ又はWebサイトが分析、調査される。

【0004】

例えば、分析者は、分析対象のWebサイトに対し、どのような他のWebサイトからアクセスされたか知ることにより、バナー広告、ディレクトリサイトへの掲載、懸賞サイトへの出品などの戦略を決定できる。

【0005】

また、分析者は、検索エンジンでどのような検索が実行され、分析対象のWebサイトがアクセスされたかなどを知ることにより、コンテンツ改善の指針を決定できる。

【0006】

Webページ又はWebサイトの分析、調査の方法として、どのWebページが、いつ、どのようにアクセスされたかを記録するWebアクセスログを用いたWebアクセスログ分析がある。

【0007】

Webアクセスログには、主に2種類ある。

【0008】

第1のWebアクセスログは、单一又は複数のWebサーバに記録されたWebアクセスログ（狭義のWebアクセスログ）である。

【0009】

第2のWebアクセスログは、一人又は複数の閲覧者による様々なWebページ又はWebサイトへのアクセスを記録したWebアクセスログである。この第2のWebアクセスログには、例えば、WWW視聴率調査会社がパネル（モニター）から収集した、各パネルによるWebページ又はWebサイトへのアクセスログがある。

【0010】

以下において、Webアクセスログとは、第1のWebアクセスログを意味するとして説明する。

【0011】

例えば、Webアクセスログ分析では、Webページ又はWebサイトにアクセスした閲覧者数（ユーザ数、訪問者数）が、時間毎、曜日毎にどのように変化するか調査される。

【0012】

また、Webアクセスログ分析では、閲覧者が使用しているブラウザ又はコンピュータのOSの種類が調査される場合もある。

【0013】

また、Web アクセスログ分析では、一つのWeb サイト内又は複数のWeb サイト間をまたがって閲覧者がどのような閲覧行動をとるかについて調査される場合もある。例えば、閲覧者がWeb ページA を閲覧し、次にWeb ページB を閲覧したことを明らかにするなどのような経路分析もWeb アクセスログ分析に含まれる。

【0014】

一方、インターネット上のWeb ページ又はWeb サイトの閲覧を補助するために、閲覧者の入力した検索キーワード（例えば、「車」「ペット」など）に応じて、この検索キーワードに関係する内容のWeb ページ（以下、「キーワード関連ページ」という）又はWeb サイト（以下、「キーワード関連サイト」という）についての情報（例えば、キーワード関連ページ又はキーワード関連サイトのURL（Uniform Resource Locators）、要約情報、リンク情報など）のリストを検索結果ページとして表示する検索システムが存在する。この検索システムは、広く利用されている。

【0015】

図45は、検索システムを用いた代表的なWeb ページの遷移の例を示す。

【0016】

検索キーワード入力画面1上で閲覧者のブラウザから検索キーワードが入力され、検索開始が指示されると、検索システムは、キーワード関連ページへのリンク2a～2cを抽出し、抽出されたキーワード関連ページへのリンク2a～2cを表示する検索結果ページ2を閲覧者に対して表示する。

【0017】

検索結果ページ2上のいずれかのリンクがクリックされると、このクリックされたリンクの先のキーワード関連ページが閲覧者のブラウザに表示される。例えば、リンク2aがクリックされると、キーワード関連ページ3が表示される。

【0018】

この検索システムを用いた場合、Web サーバのWeb アクセスログには、検索キーワードとその入力日時などを含む付加情報が記録される。

【0019】

したがって、このWebアクセスログを用いて、検索キーワードに関する分析（以下、「検索キーワード分析」という）が可能となる。

【0020】

検索キーワード分析においては、例えば、ある検索キーワードが何回入力されたか調査される。

【0021】

また、例えば、検索結果ページにおいて検索結果が実際にクリックされたか調査される。

【0022】

また、例えば、ある検索キーワードに対して、キーワード関連ページが何件抽出されたか調査される。

【0023】

特開2002-123517号（特許文献1）には、検索キーワード分析技術、商用ツール、非商用ツール、サービスが記載されている。

【0024】

【特許文献1】

特開2002-123517号公報

【0025】

【発明が解決しようとする課題】

上述したように、Webアクセスログ分析では、レファラ情報などを用いて分析対象のWebサイトをアクセスする前のWebサイトの分析、検索キーワード単位での閲覧者数の分析などは行なわれている。

【0026】

しかしながら、従来のWebアクセスログ分析では、分析者は、分析対象のWebサイトに閲覧者が満足しているか、分析対象のWebサイトにとっての成功（目的）がどの程度達成されているかまで理解することが困難である。

【0027】

例えば、分析対象のWebサイトは商品の販売数向上を目的としているとする。従来のWebアクセスログでは、検索エンジンの検索結果に基づいて分析対象

のWebサイトがアクセスされた場合にこの検索エンジンに入力された検索キーワード毎の閲覧者数を求めることはできる。

【0028】

しかし、閲覧者が多くても分析対象のWebサイトで全く商品が購入されない場合もあり、閲覧者が少なくてもほとんどの閲覧者が商品を購入する場合もある。そのため、分析者は、従来のWebアクセスログ分析の結果から、その検索キーワードを入力した閲覧者が分析対象のWebサイトに満足したか、どの検索キーワードを入力した閲覧者に対してWebサイトの目的が達成されたか判断することが困難である。

【0029】

また、例えば、従来のWebアクセスログ分析により、ある検索キーワードに対して、キーワード関連ページが何件抽出されたか調査することができる。そして、分析者は、閲覧者の希望するWebページの抽出漏れを防止する観点から、キーワード関連ページの抽出数が多いほど検索結果に対する閲覧者の満足度が高く、少ないほど満足度が低いと判断する場合がある。

【0030】

しかしながら、キーワード関連ページの抽出数は、閲覧者の満足度などを直接表すわけではない。実際には、キーワード関連ページの抽出数が少ない方が閲覧者にとってはリンクを選択しやすく、閲覧者の希望に合致し、満足度が高い場合もある。すなわち、閲覧者の希望と合致するキーワード関連ページが検索結果として閲覧者に提示される場合、閲覧者は、冗長な検索結果が提示されるよりもキーワード関連ページの抽出数が少ない程、検索結果に対して満足する。

【0031】

したがって、分析者は、従来のWebアクセスログ分析から閲覧者が検索結果に満足しているか判断することが困難である。

【0032】

また、例えば、検索結果ページ上のキーワード関連ページへのリンクが実際にクリックされたか、クリックされている場合にはどのキーワード関連ページへのリンクがクリックされたか調査すると、閲覧者によって検索結果ページ上でどの

キーワード関連ページが選択されたか知ることができる。

【0033】

しかしながら、従来のWebアクセスログ分析により、アクセスされたキーワード関連ページが本当に閲覧者の求めた内容を掲載しており閲覧者が満足したかについて、分析者が判断すること（検索結果に対して質的な評価をすること）は極めて困難である。

【0034】

本発明は、以上のような実情に鑑みてなされたもので、Webアクセスログ分析において分析者の求める分析結果を適切に提供する検索キーワード分析プログラム及びシステム並びに方法を提供することを目的とする。

【0035】

【課題を解決するための手段】

本発明を実現するにあたって講じた具体的手段について以下に説明する。

【0036】

第1の発明の検索キーワード分析プログラムは、コンピュータに、WWWにおけるアクセス要求側を示すアクセス要求側識別データとアクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録するアクセスシーケンス抽出機能と、WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データとアクセス要求側から検索システムに入力された検索キーワードとが関連付けられている場合に、検索キーワードを抽出し、検索キーワードの抽出結果を示すキーワードデータを記録手段に記録するキーワード抽出機能と、WWWにおけるアクセスの状態から検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する質的評価値を算出し、記録手段に記録されているキーワードデータの示す検索キーワードのうちアクセスシーケンスにおいて入力された検索キーワードと質的評価値との対応関係を示す関係データを記録手段に記録す

る評価値算出機能とを実現させる。

【0037】

これにより、入力された検索キーワードに基づく検索システムの検索結果に対する質的評価値を容易に把握できる。この第1の発明では、アクセスシーケンス抽出機能とキーワード抽出機能とを並列に動作させることができる。

【0038】

なお、第1の発明において、評価値算出機能は、質的評価値に加えて、検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する量的評価値を算出し、記録手段に記録されているキーワードデータの示す検索キーワードのうちアクセスシーケンスにおいて入力された検索キーワードと質的評価値と量的評価値の対応関係を示す関係データを記録手段に記録するとしてもよい。

【0039】

これにより、入力された検索キーワードに基づく検索システムの検索結果に対する質的評価値、量的評価値、質的評価値と量的評価値との関係を容易に把握できる。

【0040】

第2の発明の検索キーワード分析プログラムは、コンピュータに、WWWにおけるアクセス要求側を示すアクセス要求側識別データとアクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録するアクセスシーケンス抽出機能と、WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データとアクセス要求側から検索システムに入力された検索キーワードとが関連付けられており記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスにおいて検索キーワードが入力されている場合に、検索キーワードを抽出し、検索キーワードの抽出結果を示すキーワードデータを記録手段に記録するキーワード抽出機能と

、WWWにおけるアクセスの状態から検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する質的評価値を算出し、記録手段に記録されているキーワードデータの示す検索キーワードと質的評価値との対応関係を示す関係データを記録手段に記録する評価値算出機能とを実現させる。

【0041】

この第2の発明では、上記第1の発明と同様の効果を得ることができるが、アクセスシーケンス抽出機能とキーワード抽出機能とが直列に動作する特徴を持つ。

【0042】

なお、第2の発明において、評価値算出機能は、質的評価値に加えて、検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対応する量的評価値を算出し、記録手段に記録されているキーワードデータの示す検索キーワードと質的評価値と量的評価値の対応関係を示す関係データを記録手段に記録するとしてもよい。

【0043】

また、上記第1又は第2の発明において、Webアクセスログは、レファラ情報とアクセス先識別データとを関連付けたデータを含み、アクセスシーケンス抽出機能は、レファラ情報が検索システムを示す場合に、レファラ情報に関連付けされているアクセス先識別データの示すアクセス先へのアクセスを含むアクセスシーケンスを抽出し、キーワード抽出機能は、レファラ情報が検索システムを示す場合に、レファラ情報に含まれている検索キーワードを抽出し、評価値算出機能は、アクセスシーケンスに対する検索キーワード入力後の質的評価値を算出するとしてもよい。

【0044】

このように、レファラ情報を用いることでにより、アクセス先で記録されたWebアクセスログのみからであってもサイト外の検索システムに入力された検索

キーワードを分析することができる。

【0045】

第3の発明の検索キーワード分析プログラムは、コンピュータに、WWWにおけるアクセス要求側を示すアクセス要求側識別データとアクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとレファラ情報とを関連付けてWebアクセスログに基づいて、レファラ情報がWWWの検索システムを示す場合にレファラ情報に関連付けされているアクセス先識別データの示すアクセス先へのアクセスを含み同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出し、アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを記録手段に記録するアクセスシーケンス抽出機能と、レファラ情報が検索システムを示す場合に、レファラ情報に含まれている検索キーワードを抽出し、検索キーワードの抽出結果を示すキーワードデータを記録手段に記録するキーワード抽出機能と、WWWにおけるアクセスの状態から検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対する検索キーワード入力後の質的評価値を算出し、記録手段に記録されているキーワードデータの示す検索キーワードのうちアクセスシーケンスに対応する検索キーワードとアクセスシーケンスに対応する質的評価値との対応関係を示す関係データを前記記録手段に記録する評価値算出機能とを実現させる。

【0046】

なお、第3の発明において、評価値算出機能は、質的評価値に加えて、検索システムの検索結果に対する量的評価値を算出する基準を定義した量的評価規則に基づいて、記録手段に記録されたアクセスシーケンスデータの示すアクセスシーケンスに対する検索キーワード入力後の量的評価値を算出し、記録手段に記録されているキーワードデータの示す検索キーワードと質的評価値と量的評価値の対応関係を示す関係データを記録手段に記録するとしてもよい。

【0047】

上記第1乃至第3の発明の検索キーワード分析プログラムにおいて、検索キーワードと質的評価値と量的評価値の対応関係を示す関係データが求められる場合

、コンピュータに、記録手段に記録された関係データにしたがって、質的評価値と量的評価値とに基づくグラフを描画するグラフ化機能を実現させるためのプログラム、を付加してもよい。これにより、分析結果の把握が一層容易になる。

【0048】

上記第1乃至第3の発明の検索キーワード分析プログラムにおいて、検索キーワードと質的評価値と量的評価値とが求められる場合、量的評価値は、検索キーワードの入力数が多いほど値が大きくなる検索結果のニーズ値であり、量的評価規則は、検索キーワードの入力数が多いほどニーズ値が増加する規則であるとしてもよい。

【0049】

上記第1乃至第3の発明の検索キーワード分析プログラムにおいて、質的評価規則は、分析候補の検索キーワードと分析候補の検索キーワードの入力後にアクセスされた場合に成功度を増加するアクセス先識別データとを関連付けたデータと記録手段に記録されたキーワードデータとに基づいて、アクセスシーケンスにおいて分析候補の検索キーワードが入力されておりその入力後分析候補の検索キーワードと関連付けされているアクセス先識別データの示すアクセス先がアクセスされている場合に、アクセスシーケンスに対する成功度を増加する規則であり、質的評価値は、成功度であるとしてもよい。

【0050】

この場合において、質的評価規則は、分析候補の検索キーワードが入力されてから分析候補の検索キーワードと関連付けされているアクセス先識別データの示すアクセス先がアクセスされるまでのアクセス数が少ないほど加えられる成功度を大きくするとしてもよい。

【0051】

上記第1乃至第3の発明の検索キーワード分析プログラムにおいて、Webアクセスログは、WWWにおけるアクセス毎に、アクセス要求側識別データとアクセス先識別データとアクセス時刻を示すデータとを関係付けており、質的評価規則は、アクセスシーケンスにおける検索キーワード入力後のアクセスから最後のアクセスまでの時間間隔が長いほどアクセス要求側の満足度が増加する規則であ

り、質的評価値は、満足度であるとしてもよい。

【0052】

第4の発明の検索キーワード分析プログラムは、コンピュータに、WWWにおけるアクセス要求側を示すアクセス要求側識別データとアクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスのうち検索結果ページへのアクセスから定まる始端と、始端と同一のアクセス要求側による連続的なアクセスのうち始端以降のアクセスから定まる終端とに基づくサブアクセスシーケンスを抽出し、サブアクセスシーケンスの抽出結果を示すサブアクセスシーケンスデータを記録手段に記録するサブアクセスシーケンス抽出機能と、WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データとアクセス要求側から検索システムに入力された検索キーワードとが関連付けされている場合に、検索キーワードを抽出し、検索キーワードの抽出結果を示すキーワードデータを記録手段に記録するキーワード抽出機能と、WWWにおけるアクセスの状態から検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、記録手段に記録されたサブアクセスシーケンスデータの示すサブアクセスシーケンスに対応する質的評価値を算出し、記録手段に記録されているキーワードデータの示す検索キーワードのうちサブアクセスシーケンスに対応する検索キーワードと質的評価値との対応関係を示す関係データを記録手段に記録する評価値算出機能とを実現させる。

【0053】

第5の発明の検索キーワード分析プログラムは、コンピュータに、WWWにおけるアクセス要求側を示すアクセス要求側識別データとアクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスから、検索結果ページへのアクセスから定まる始端と次の始端とを求めるとともに次の始端の直前のアクセスに基づいて始端に対応する終端を求めてサブアクセスシーケンスを抽出し、サブアクセスシーケンスの抽出結果を示すサブアクセスシーケンスデータを記録手段に記録するサブアクセスシーケンス抽出機能と、Webアクセスロ

グにおいてWWWの検索システムを示すアクセス先識別データとアクセス要求側から検索システムに入力された検索キーワードとが関連付けされている場合に、検索キーワードを抽出し、検索キーワードの抽出結果を示すキーワードデータを記録手段に記録するキーワード抽出機能と、WWWにおけるアクセスの状態から検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、記録手段に記録されたサブアクセスシーケンスデータの示すサブアクセスシーケンスに対応する質的評価値を算出し、記録手段に記録されているキーワードデータの示す検索キーワードのうちサブアクセスシーケンスに対応する検索キーワードと質的評価値との対応関係を示す関係データを記録手段に記録する評価値算出機能とを実現させる。

【0054】

第6の発明の検索キーワード分析プログラムは、コンピュータに、コンピュータに、WWWにおけるアクセス要求側を示すアクセス要求側識別データとアクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebアクセスログに基づいて、同一のアクセス要求側による連続的なアクセスから検索結果ページへのアクセスから定まる始端と次の始端とが求められた場合、次の始端の直前のアクセスを始端に対応する終端とし、始端が求められ次の始端が求められなかった場合、連続的なアクセスの終端を始端に対応する終端とし、始端と終端とにに基づいてサブアクセスシーケンスを抽出し、サブアクセスシーケンスの抽出結果を示すサブアクセスシーケンスデータを記録手段に記録するサブアクセスシーケンス抽出機能と、WebアクセスログにおいてWWWの検索システムを示すアクセス先識別データとアクセス要求側から検索システムに入力された検索キーワードとが関連付けられている場合に、検索キーワードを抽出し、検索キーワードの抽出結果を示すキーワードデータを記録手段に記録するキーワード抽出機能と、WWWにおけるアクセスの状態から検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいて、記録手段に記録されたサブアクセスシーケンスデータの示すサブアクセスシーケンスに対応する質的評価値を算出し、記録手段に記録されているキーワードデータの示す検索キーワードのうちサブアクセスシーケンスに対応する検索キーワードと質

的評価値との対応関係を示す関係データを記録手段に記録する評価値算出機能とを実現させる。

【0055】

上記第1乃至第3の発明の検索キーワード分析プログラムにおいて、検索キーワードと質的評価値と量的評価値の対応関係を示す関係データが求められる場合、Webアクセスログは、アクセス要求側識別データとアクセス先識別データとアクセスされた時間を示す時間データとを関係付けたデータであり、評価値算出機能は、時間データに基づいて複数の期間毎の関係データを求め、コンピュータに、複数の期間毎の関係データに基づいて、検索キーワードについて質的評価値と量的評価値との遷移状況を表すトレンドデータを求め、記録手段に記録するトレンド分析機能を実現させるとしてもよい。

【0056】

これにより、WWWにおける分析対象の部分をアクセスするために検索システムに入力される検索キーワードの質的評価値及び量的評価値の変化が時系列に求められる。

【0057】

上記第1乃至第3の発明の検索キーワード分析プログラムにおいて、検索キーワードと質的評価値と量的評価値の対応関係を示す関係データが求められる場合、Webアクセスログは、アクセス要求側識別データとアクセス先識別データとアクセスされた時間を示す時間データとを関係付けたデータであり、評価値算出機能は、時間データに基づいて複数の期間毎の前記関係データを求め、コンピュータに、検索キーワードと検索キーワードの属するカテゴリとを関係付けたカテゴリデータと複数の期間毎の関係データとにに基づいて複数の期間毎にカテゴリについて質的評価値と量的評価値とを算出し、記録手段に記録するカテゴリ形成機能と、カテゴリについて質的評価値と量的評価値との遷移状況を表すトレンドデータを求め、記録手段に記録するトレンド分析機能とを実現させるとしてもよい。

【0058】

これにより、検索キーワードを統合して分析者の求めるカテゴリについて分析

を行うことができる。

【0059】

上記の検索キーワード分析プログラム又はこのプログラムを記録した記録媒体を用いることによって、上記各機能を有していないコンピュータ、コンピュータシステム、サーバ、クライアントに対しても、簡単にこれらの機能を付加することができる。

【0060】

本発明で実現される機能と同様の動作を実現する手段を検索キーワード分析システムに具備することにより、同様の作用効果を得ることができる。

【0061】

本発明の検索キーワード分析プログラム又はこのプログラムを記録した記録媒体をコンピュータに適用することによって、入力された検索キーワードに基づく検索システムの検索結果に対する質的評価値を容易に把握するための検索キーワード分析方法がコンピュータ上で実施される。

【0062】

【発明の実施の形態】

以下、図面を参照しながら本発明の実施の形態について説明する。なお、Webサイトとは少なくとも1つのWebページを含むWebページのグループを指す。

【0063】

以下の各図において同一の部分については同一の符号を付してその説明を省略し、異なる部分についてのみ詳しく説明する。

【0064】

(第1の実施の形態)

本実施の形態においては、検索システムによる検索結果を表示する検索結果ページからの閲覧者の閲覧行動（2つ以上のページ遷移を含む）を分析することにより、閲覧者から入力された検索キーワードに基づく検索結果に対する質的評価値を高精度に求める検索キーワード分析システムについて説明する。

【0065】

上述したように、検索システムは、閲覧者が入力した検索キーワードに応じて、この検索キーワードと関連があると判断されたキーワード関連ページについての情報を抽出し、リスト化し、閲覧者に表示する。

【0066】

閲覧者は、検索システムに対して自己の意思に基づいて自由に検索キーワードを入力する。

【0067】

このため、各種の検索キーワードの入力頻度は、各種の検索キーワード間で一定の割合となることはなく、ある程度隔たりが発生する。

【0068】

したがって、顧客サイトをアクセスするために入力された検索キーワードによる検索結果を量的に評価する量的評価値は、検索キーワード毎に差が生じる。

【0069】

量的評価値の例としては、各検索キーワードについて検索結果のニーズの量、数、多さを表すニーズ値があり、例えば検索キーワードの入力回数をニーズ値とすることができます。

【0070】

その他にも、量的評価値の例としては、上述した検索キーワードについて抽出されたキーワード関連ページの件数、検索結果ページのクリック回数などがある。

【0071】

その一方で、検索キーワードに基づいて検索されたキーワード関連ページが必ずしも閲覧者の求めている内容とは限らない。

【0072】

したがって、顧客サイトをアクセスするために入力された検索キーワードによる検索結果を質的に評価する質的評価値を求める意義は大きい。

【0073】

質的評価値の例としては、検索キーワードによる検索結果がどの程度効果的に機能したかを表す度合いである効果度がある。さらに、この効果度の具体例とし

では、検索システムを利用した閲覧者が、その検索キーワードに対して検索システムが提示した情報サービスにどの程度満足したかを表す満足度がある。また、効果度の他の具体例として、Webサイト運営者がアクセスして欲しいと考えているWebページにどの程度自サイトの閲覧者がアクセスしているかを表す成功度がある。

【0074】

質的評価値として、例えば、閲覧者が検索キーワードによる検索結果をどの程度重視しているかを表す重視度もある。

【0075】

なお、質的評価値についても、検索システムの性能、リストとして表示されたキーワード関連ページの質、内容などに応じて、検索キーワード毎に差が生じる。

【0076】

質的評価値を調査することは、Webページ又はWebサイトの運営者にとって、自己の運営しているWebページ又はWebサイト、検索システムが閲覧者を例えどどの程度満足させているか把握するために有効であり、Webページ又はWebサイトのデザインやコンテンツ、検索システムの改善に有効である。

【0077】

なお、本実施の形態においては、質的評価値のうち満足度を求める場合について説明するが、満足度ではない他の質的評価値を求める場合も同様である。

【0078】

図1は、本実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図である。

【0079】

図2は、顧客サイトの構成の一例を示すブロック図である。

【0080】

記録装置4は、調査対象Webサイトの運営者である顧客5によって運営される顧客サイト6のWebアクセスログ7を記録する。

【0081】

顧客サイト6は、サイト内検索システム6aと、一以上のWebページP₁～P_nを持つ。

【0082】

サイト内検索システム6aは、ある特定のサイト内のWebページについての検索を行う。この図1の例では、サイト内検索システム6aは、顧客サイト6のWebページP₁～P_nについての検索を行う。

【0083】

一般的な閲覧者8のうち、顧客サイト6をアクセスした閲覧者を顧客サイト閲覧者8aとする。

【0084】

表1に、Webアクセスログ7に含まれる情報の一例を示す。

【0085】

【表1】

表1 Webアクセスログに含まれる情報（顧客サイト閲覧者）

ホスト名	User ID	日付	時刻	リクエスト	終了コード	転送バイト数	レフアラ情報	ユーザエージェント
pc001.000.co.jp	—	2002/2/14	18:56:14	/index.html	200	400	—	MOOiiia/4.0 (compatible:MOIE...)
ws20.△△△.co.jp	—	2002/2/14	18:57:30	/cgi/login.cgi	200	700	—	MOOiiia/4.7 [△△5LLL](X11...)
pc001.000.co.jp	—	2002/2/14	18:58:01	/products/index.html	200	2400	—	w△m/0.3+m□e-p21-26+mOe-1.5.2
pc001.000.co.jp	—	2002/2/14	18:58:40	/products/info/products/p001.pdf	200	3000	—	MOOiiia/4.7 [△△5LLL](X11...)
ws20.△△△.co.jp	—	2002/2/14	18:59:10	/cgi/order.cgi	200	700	—	MOOiiia/4.0 (compatible:MOIE...)

【0086】

W e b アクセスログ 7 には、顧客サイト閲覧者 8 a による顧客サイト 6 のアクセス毎に、ホスト名又は I P アドレス、顧客サイト閲覧者 8 a が顧客サイト 6 にアクセスした日付と時刻を示す情報、顧客サイト閲覧者 8 a の発信したリクエスト (U R L) 、レファラ情報、アクセスした顧客サイト閲覧者 8 a の I D (User ID) 、などから構成されるレコードが含まれる。

【0087】

なお、レファラ情報とは、ハイパーリンクを使ってあるW e b ページがアクセスされた場合に、このアクセスされたW e b ページがどのW e b ページからリンクを辿ってアクセスされたかを示す情報である。表 1 においては、顧客サイト閲覧者 8 a の I D とレファラ情報については、具体的な値の記述を省略している。

【0088】

本実施の形態に係る検索キーワード分析システム 9 は、記録媒体 1 0 に記録されている検索キーワード分析プログラム 1 1 を読み出して実行する。

【0089】

検索キーワード分析プログラム 1 1 は、システム 9 に読み込まれ、実行されると、ログ前処理機能 1 2 、アクセスシーケンス抽出機能 1 3 、キーワード抽出機能 1 4 、評価値算出機能 3 7 を実現する。

【0090】

また、評価値算出機能 3 7 は、満足度算出機能 1 5 、満足度・キーワード関連付け機能 1 6 を含む。

【0091】

上記各機能 1 2 ～ 1 6 によって作成されたデータは、記録装置 3 8 に記録される。各機能 1 2 ～ 1 6 は、記録装置 3 8 に記録されたデータを読み込み可能である。記録装置 3 8 としては、例えば内部メモリ等が適用される。すなわち、各機能 1 2 ～ 1 6 間での各種データの受け渡しは、記録装置 3 8 を経由して行われてもよい。この場合、データの提供側は、提供するデータを記録装置 3 8 に記録し、データの受付側は、記録装置 3 8 からデータを読み出す。

【0092】

ログ前処理機能 1 2 は、記録装置 4 に記録されたW e b アクセスログ 7 から、

検索キーワード分析システム9に必要な情報を入力する。

【0093】

そして、ログ前処理機能12は、入力した情報に基づいて、各ページアクセスに対して重複のないアクセス識別子(ViewID)を付与する。

【0094】

表2に、ログ前処理機能12によって生成された前処理データの一例を示す。

【0095】

【表2】

表2 前処理データ

ViewID	ホスト名	User ID	日付	時刻	リクエスト	レファラ情報	ユーザエージェント
View1	pc001.000.co.jp	—	2002/2/14	18:56:14	/index.html	—	MOOiiia/4.0 (compatible:MOIE...)
View2	ws20.△△△.co.jp	—	2002/2/14	18:57:30	/cgi/login.cgi	—	MOOiiia/4.7 [△△5LLL](X11...)
View3	pc001.000.co.jp	—	2002/2/14	18:58:01	/products/index.html	—	W△m/0.3+m□e-p21-26+mOe-1.5.2
View4	pc001.000.co.jp	—	2002/2/14	18:58:40	/products/info/p001.pdf	—	MOOiiia/4.7 [△△5LLL](X11...)
View5	ws20.△△△.co.jp	—	2002/2/14	18:59:10	/cgi/order.cgi	—	MOOiiia/4.0 (compatible:MOIE...)

【0096】

アクセスシーケンス抽出機能13は、前処理データに基づいて、顧客サイト閲

観者 8 a によってアクセスされた一連の Web ページの列（以下、「アクセスシーケンス」という）を抽出し、各アクセスシーケンスに重複のないアクセスシーケンス識別子を付与し、アクセスシーケンス抽出データを生成する。

【0097】

アクセスシーケンスとは、同一又は同一として扱われる閲覧者、すなわち同一のアクセス要求側による連続的なアクセスをいう。

【0098】

アクセスシーケンス抽出機能 13 は、ホスト名又は IP アドレスとユーザエージェントとの組み合わせが一致する場合に同一のアクセス要求側として扱い、ホスト名又は IP アドレスとユーザエージェントの各組み合わせに対してアクセス要求側識別子を付与し、同一のアクセス要求側から連続的にアクセスされた一以上の Web ページをアクセスシーケンスとして切り出す。

【0099】

アクセスシーケンス抽出機能 13 は、所定のアクセスシーケンス抽出規則にしたがって、一連のページアクセスからアクセスシーケンスを切り出し、アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを作成する。

【0100】

例えば、アクセスシーケンス抽出規則として、一般的な Web アクセスログ分析で用いられているセッションを抽出する規則を適用可能である。具体的には、ある閲覧者による Web ページのアクセスから次の Web ページへのアクセスの間に 30 分以上の時間が経過した場合（次の Web ページへの遷移がなかった場合を含む）、その間をセッションの境界（もしくはセッションの終端）とし、境界と境界との間をセッションとして抽出する。そして、このセッションをアクセスシーケンスとして用いる。

【0101】

なお、アクセスシーケンスとアクセスシーケンスとの間の境界を判定するための経過時間は、上記の 30 分に限定されるものではなく、自由に設定可能である。すなわち、アクセスシーケンス抽出規則は、複数のアクセスが同一のアクセス要求側から所定の時間間隔内でアクセスされている場合に複数のアクセスが同一

のアクセスシーケンスに含まれると判断する規則としてもよい。

【0102】

また、アクセスシーケンスを抽出する場合において、Webアクセスログ7に記録された閲覧者のID（UserID）又はCookie情報に基づいて、アクセス要求側の同一性を判断してもよい。

【0103】

表3に、アクセスシーケンスデータの一例を示す。

【0104】

【表3】

表3 アクセスシーケンスデータ

ViewID	VisitID	VisitorID	日付	時刻	リクエスト
View1	Visit1	Visitor1	2002/2/14	18:56:14	/AAA.html
View2	Visit2	Visitor2	2002/2/14	18:57:30	/BBB.cgi
View3	Visit1	Visitor1	2002/2/14	18:58:01	/CCC.html
View4	Visit2	Visitor2	2002/2/14	18:58:40	/GGG/DDD.cgi
View5	Visit3	Visitor1	2002/2/14	19:42:10	/EEE.html

【0105】

表3のアクセスシーケンスデータには、各アクセスシーケンスを識別するためのアクセスシーケンス識別子（VisitID）、ホスト名とユーザエージェントの組み合わせを識別するアクセス要求側識別子（VisitorID）などから構成されるコードが含まれる。

【0106】

キーワード抽出機能14は、ログ前処理機能12によって生成された前処理データの中から、顧客サイト閲覧者8aがサイト内検索システム6aにおいて入力した検索キーワードを抽出し、検索キーワードの抽出結果を示すキーワードデータを作成する。

【0107】

表4に、Webアクセスログ7における検索システムへのリクエストの例を示す。

【0108】

【表4】

表4 検索システムへのリクエスト

例1	http://www.○○○.co.jp/search.cgi?q=xxxx
例2	http://www.○○○.co.jp/search.cgi?q=yyyy&page=10
例3	http://www.○○○.co.jp/search.cgi?q=%E3%82%8A%E3%82%93%E3%81%94

【0109】

表4における例1は、検索キーワード「xxxx」を含んだ一番基本的な検索システムへのリクエストの例である。

【0110】

例2では、検索キーワード「yyyy」と検索結果の10ページ目とがリクエストされている。

【0111】

例3では、検索キーワード「りんご」を示すコードがリクエストに含まれている。

【0112】

Webアクセスログ7の中から検索キーワードを抽出するためには、検索システムの場所及び検索キーワードを示すパラメータを表すクエリーパターンに関する情報を用いる。

【0113】

表5に、検索システムの場所とクエリーパターンに関する情報の一例を示す。

【0114】

【表5】

表5 検索システムの場所とクエリーパターン

検索システムの場所	クエリーパターン
http://www.○○○.co.jp/search.cgi	q

【0115】

すなわち、キーワード抽出機能14は、前処理データから検索システムの場所に関する情報に基づいて、検索システムを利用したアクセスを抽出する。

【0116】

次に、キーワード抽出機能14は、クエリーパターンに関する情報に基づいて、検索キーワードを抽出する。

【0117】

表6に、上記表4、5の例に基づいて、キーワード抽出機能14によって抽出されたパラメータの例を示す。

【0118】

【表6】

表6 抽出されたパラメータ

	パラメータ
例1	xxxx
例2	yyyy
例3	%E3%82%8A%E3%82%93%E3%81%94

【0119】

抽出されたパラメータの中には、この表6の例3に示すようにエンコードされたパラメータも存在する。

【0120】

キーワード抽出機能14は、エンコードされたパラメータについては判読可能な文字列へのデコードを行い、キーワードデータを生成する。

【0121】

表7に、キーワードデータの例を示す。

【0122】

【表7】

表7 キーワードデータ

ViewID	検索キーワード
View6	りんご
View21	apple
View43	banana
View56	砂糖
View81	カレイ

【0123】

表7のキーワードデータには、各アクセス識別子とそのアクセス識別子に対応するアクセスにおいて入力された検索キーワードとから構成されるレコードが含まれている。

【0124】

満足度算出機能15は、アクセスシーケンス抽出機能13によって生成されたアクセスシーケンスデータに基づいて、閲覧者の満足度を算出する。

【0125】

満足度を算出するために、満足度をどのように算出するかを表す質的評価規則が定義される。

【0126】

例えば、質的評価規則として、サイト内検索システム6aに検索キーワードを入力した後の顧客サイト閲覧者8aによる顧客サイト6の閲覧時間（以下、「検索後閲覧時間」という）が長いほど満足度が高くなる規則を定義する。

【0127】

表8に、検索後閲覧時間と満足度とを関係付けた例を示す。

【0128】

【表8】

表8 検索後閲覧時間と満足度

検索後閲覧時間	満足度
閲覧なし	0
1ページのみの閲覧	0
1分以下	10
1-2分	20
2-3分	30
3-5分	40
5-7分	50
7-10分	60
10-15分	70
15-20分	80
20-25分	90
30分以上	100

【0129】

検索後閲覧時間は、満足度を求めるアクセスシーケンスについて最後のアクセス時刻から最初のキーワード関連ページへのアクセス時刻を引いて得られる時間間隔とする。ただし、検索結果ページからキーワード関連ページに遷移することなくアクセスシーケンスが終わっていた場合には検索後閲覧時間を「閲覧なし」とする。また、検索結果ページからキーワード関連ページに遷移して、その後他のサイト内のWebページを閲覧することなくアクセスシーケンスが終わっている場合には、検索後閲覧時間を「1ページのみの閲覧」とする。

【0130】

すなわち、満足度算出機能15は、アクセスシーケンスデータを読み込み、アクセスシーケンスごとに検索後閲覧時間を計算する。

【0131】

次に、満足度算出機能15は、上記表8のような検索後閲覧時間と満足度の対応データを読み込み、アクセスシーケンスごとの満足度を算出する。

【0132】

表9に、満足度算出機能15によって生成された満足度算出データの例を示す。

【0133】

【表9】

表9 満足度算出データ

ViewID	VisitID	満足度
View1	Visit1	0
View2	Visit2	20
View3	Visit1	0
View4	Visit2	20
View5	Visit3	10

【0134】

表9の満足度算出データには、各アクセス識別子とそのアクセスの属するアクセスシーケンス識別子に対応する満足度とから構成されるレコードが含まれている。

【0135】

満足度・キーワード関連付け機能16は、満足度算出機能15で生成された満足度算出データと、キーワード抽出機能14で生成されたキーワードデータを統合し、検索キーワードごとの満足度を算出し、満足度と検索キーワードとを関係付け、関係データを作成する。

【0136】

満足度・キーワード関連付け機能16は、まず、満足度算出データとキーワードデータを読み込む。

【0137】

次に、満足度・キーワード関連付け機能16は、アクセス識別子に基づいて、満足度算出データとキーワードデータを統合する。

【0138】

表10に、満足度算出データとキーワードデータの統合結果を示す。

【0139】

【表10】

表10 満足度算出データとキーワードデータの統合結果

ViewID	満足度	検索キーワード
View6	10	りんご
View21	0	apple
View83	30	りんご
View121	20	砂糖
View137	0	apple

【0140】

そして、満足度・キーワード関連付け機能16は、検索キーワードごとに満足度の平均を求め、関係データを作成する。

【0141】

表11に、関係データの例を示す。

【0142】

【表11】

表11 関係データ

検索キーワード	満足度
りんご	23.4
apple	17.8
砂糖	2.4
肉野菜	2.4
蜂蜜	11.6

【0143】

図3は、本実施の形態に係る検索キーワード分析システム9の動作の一例を示すフローチャートである。

【0144】

ステップS1において、ログ前処理機能12は、Webアクセスログ7に含まれる情報から必要な情報を入力し、前処理データを作成する。

【0145】

ステップS2において、アクセスシーケンス抽出機能13は、前処理データとアクセスシーケンス抽出規則とに基づいて、アクセスシーケンスを抽出し、アクセスシーケンスデータを作成する。

【0146】

ステップS3において、満足度算出機能15は、アクセスシーケンスデータと満足度算出規則とに基づいて、アクセスシーケンス毎の満足度を算出し、満足度算出データを作成する。

【0147】

ステップS4において、キーワード抽出機能14は、前処理データに基づいて、サイト内検索システム6aに入力された検索キーワードを抽出し、キーワードデータを作成する。

【0148】

ステップS5において、満足度・キーワード関連付け機能16は、満足度算出データと、キーワードデータに基づいて、検索キーワードごとの満足度を算出し、関係データを作成する。

【0149】

なお、ステップS4は、ステップS2及びステップS3の前に実行されてもよく、後に実行されてもよく、並列に実行されてもよい。

【0150】

図4は、本実施の形態に係るキーワード抽出機能14の動作の一例を示すフローチャートである。

【0151】

ステップT1において、キーワード抽出機能14は、検索システムの場所とクエリーパターンに関する情報を読み込む。

【0152】

ステップT2において、キーワード抽出機能14は、ログ前処理機能12によって前処理されたWebアクセスログ7から一つのレコードを読み込む。

【0153】

ステップT3において、キーワード抽出機能14は、読み込んだレコードに含まれているリクエストが検索システムの場所を含むか判定する。

【0154】

リクエストが検索システムの場所を含むと判定された場合、ステップT4において、キーワード抽出機能14は、検索システムの場所に対応するクエリーパターンに基づいて、リクエストからパラメータを抽出する。

【0155】

リクエストが検索システムの場所を含むと判定されなかった場合、キーワード抽出機能14は、ステップT2以降の処理を繰り返す。

【0156】

ステップT5において、キーワード抽出機能14は、抽出したパラメータに対してデコードが必要か否か判定する。

【0157】

デコードが必要な場合、ステップT6において、キーワード抽出機能14は、抽出したパラメータをデコードする。

【0158】

ステップT 7において、キーワード抽出機能14は、ログ前処理機能12によって前処理されたWebアクセスログ7から終端のレコードを読み込んだか否か判定する。

【0159】

終端のレコードが読み込まれていないと判定された場合、キーワード抽出機能14は、ステップT 2以降の処理を繰り返す。

【0160】

終端のレコードが読み込まれたと判定された場合、ステップT 8において、キーワード抽出機能14は、抽出されたパラメータを検索キーワードとし、キーワードデータを作成する。

【0161】

図5は、本実施の形態に係る満足度算出機能15の動作の一例を示すフローチャートである。

【0162】

ステップU 1において、満足度算出機能15は、アクセスシーケンス抽出機能13によって生成されたアクセスシーケンスデータを読み込む。

【0163】

ステップU 2において、満足度算出機能15は、アクセスシーケンス毎に、検索後閲覧時間を計算する。

【0164】

ステップU 3において、満足度算出機能15は、検索後閲覧時間と満足度とを関係付けたデータを読み込む。

【0165】

ステップU 4において、満足度算出機能15は、アクセスシーケンス毎に、検索後閲覧時間に対応する満足度を求める。

【0166】

ステップU 5において、満足度算出機能15は、アクセスシーケンス毎の満足度を示す満足度算出データを作成する。

【0167】

なお、ステップU3は、ステップU1及びステップU2の前に実行されてもよく、並列に実行されてもよい。

【0168】

以上説明した本実施の形態においては、検索キーワードについての分析結果として、顧客サイト閲覧者8aがその検索キーワードによる検索結果についてどれほど満足したか求めることができ、検索キーワードに対する検索結果について量的評価値のみではなく、十分な精度で質的評価値を分析することができる。

【0169】

本実施の形態により、例えば、顧客5は、検索結果に対して顧客サイト閲覧者8aがどの程度満足しているかを定量的に知ることができる。例えば、上記表11の結果より、検索キーワード「りんご」による検索結果については顧客サイト閲覧者8aの満足度が高いのに対し、検索キーワード「砂糖」又は検索キーワード「肉野菜」による検索結果については顧客サイト閲覧者8aの満足度が低いことが分かる。

【0170】

なお、質的評価値としては、満足度等に代表される効果度の他に、例えば重視度や利便度も考えられる。

【0171】

重視度とは、Webサイトの閲覧者が検索キーワードに関連する事柄について、どの程度重視しているかを表す指標である。

【0172】

利便度とは、Webサイトの閲覧者が検索キーワードに関連する事柄について、どの程度容易に閲覧やWebサイトへの書き込み等の操作ができたかを表す指標である。

【0173】

重視度や利便度も、満足度と同様に検索後閲覧時間によって求めることが可能である。

【0174】

検索後閲覧時間を満足度、重視度、利便度のいずれを算出するのに利用するか

は、分析対象となるWebサイトや閲覧者の性質、分析者の観点によって適宜選択できる。

【0175】

例えば、商品を販売しているECサイトにおいて、他にも類似の競合サイトが多く存在するケースでは、顧客は検索キーワードに関連したコンテンツが顧客の期待にそわなかつた場合に、比較的短時間で閲覧をあきらめてサイトを去り、別のサイトを閲覧することが考えられる。このようなケースでは、検索後閲覧時間は満足度としての利用が適切である。例えば、「蜂蜜」という検索キーワードに対する検索後閲覧時間が短いために、「蜂蜜に関する満足度が低い」という分析結果が得られた場合、この結果は「蜂蜜」に関連するコンテンツを改良して、満足度を向上させるような対策を打つ必要性が高い、という判断をする根拠として有効である。

【0176】

一方、他に類似の競合サイトが少ない場合や、提供しているコンテンツが閲覧者の嗜好を満たすような内容というよりはむしろ合理的必要性に応じて閲覧している場合や、その検索キーワードに関する情報をできるだけ当該Webサイトから得たいと思っている閲覧者が多数であるような場合においては、閲覧者が検索キーワードに関連する内容を重視している度合いに応じて、長時間サイトに滞在して、その情報を閲覧すると考えられる。したがって、検索後閲覧時間を重視度の算出に利用することが適切な場合がある。この場合に、ある検索キーワードに対する検索後閲覧時間が短いために、それに関する重視度が低いという分析結果が得られた場合、この結果はそのコンテンツについては以後それほど改良のための対策を打つ必要性が低い、という判断をする根拠として有効である。

【0177】

さらに、利便度においては、例えば提供しているコンテンツをどれだけ閲覧者が容易に理解したか、情報の入力がどの程度容易であったかを評価する指標として有効である。

【0178】

例えば検索キーワードに応じてアンケート内容を変えているWebサイトにお

いて、検索後閲覧時間が長いということは、アンケートの返答の入力に手間取っていると判断され、利便度が低いと判断される。

【0179】

このように利便度の低いコンテンツがあれば、利便度を向上させるべく、対策を打つ必要性が高い、という判断をする根拠として有効である。

【0180】

（第2の実施の形態）

本実施の形態においては、上記第1の実施の形態に係る検索キーワード分析システムの変形例について説明する。なお、本実施の形態においては、質的評価値のうち満足度を求め、量的評価値のうちニーズ値を求める場合について説明するが、他の質的評価値及び量的評価値を求める場合も同様である。

【0181】

図6は、本実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図である。

【0182】

本実施の形態に係る検索キーワード分析システム18は、記録媒体19に記録されている検索キーワード分析プログラム20を読み出して実行する。

【0183】

検索キーワード分析プログラム20は、システム18に読み込まれ、実行されると、ログ前処理機能12、アクセスシーケンス抽出機能13、キーワード抽出機能14、評価値算出機能39を実現する。

【0184】

また、評価値算出機能39は、満足度算出機能15、ニーズ値算出機能21、満足度・ニーズ値関連付け機能22を含む。

【0185】

ニーズ値算出機能21、満足度・ニーズ値関連付け機能22によって作成されたデータは、記録装置38に記録される。ニーズ値算出機能21、満足度・ニーズ値関連付け機能22は、記録装置38に記録されたデータを読み込み可能である。

【0186】

ニーズ値算出機能21は、キーワード抽出機能14によって生成されたキーワードデータに基づいて、検索キーワードごとのニーズ値を算出する。

【0187】

ニーズ値を算出するために、ニーズ値をどのように算出するかを表す量的評価規則が定義される。

【0188】

例えば、量的評価規則として、検索キーワードの入力総数に対する各検索キーワードの占める割合を百分率で表した値をニーズ値と定義する。

【0189】

ニーズ値算出機能21は、キーワード抽出機能14によって生成されたキーワードデータを読み込み、検索キーワードごとの入力個数をカウントする。

【0190】

次に、ニーズ値算出機能21は、検索キーワードごとの入力個数から全検索キーワードの入力個数の合計を算出する。

【0191】

次に、ニーズ値算出機能21は、全検索キーワードの入力個数と検索キーワードごとの入力個数に基づいて、全検索キーワードの入力に対する各検索キーワードの占める割合を百分率で表す。

【0192】

そして、ニーズ値算出機能21は、検索キーワードごとのニーズ値を表すニーズ値算出データを作成する。

【0193】

表12に、ニーズ値算出データの一例を示す。

【0194】

【表12】

表12 ニーズ値算出データ

検索キーワード	ニーズ値
りんご	21.5
apple	14.6
砂糖	3.6
肉野菜	17.6
蜂蜜	2.7

【0195】

満足度・ニーズ値関連付け機能22は、満足度算出機能15によって生成された満足度算出データとニーズ値算出機能21によって生成されたニーズ値算出データを読み込む。

【0196】

そして、満足度・ニーズ値関連付け機能22は、満足度算出データとニーズ値算出データを検索キーワード毎に統合し、検索キーワードごとのニーズ値、満足度を表す関係データを作成する。

【0197】

表13に、関係データの一例を示す。

【0198】

【表13】

表13 関係データ

検索キーワード	ニーズ値	満足度
りんご	21.5	23.4
apple	14.6	17.8
砂糖	3.6	2.4
肉野菜	17.6	2.4
蜂蜜	2.7	11.6

【0199】

以上説明した本実施の形態において、顧客5は、質的評価値（例えば満足度）と量的評価値（例えばニーズ値）とを参照、比較しつつ、早急に対策を取る必要がある検索キーワードについて知ることができる。

【0200】

例えば、顧客5は、上記表13の例を参照した場合に、検索キーワード「砂糖」と「肉野菜」は同程度の満足度であり他の検索キーワードに比べて満足度が低いが、「肉野菜」は「砂糖」に比べてニーズ値が高いと解釈できる。

【0201】

この結果、顧客5は、検索キーワード「砂糖」について満足度が低いあまり検索のニーズがなく（本実施形態の場合、検索キーワード「砂糖」による検索結果ページの閲覧者が少なく）、比較的影響力が低いと判断できる。

【0202】

これに対し、顧客5は、検索キーワード「肉野菜」は「砂糖」と同じ程度で満足度が低いが検索のニーズがあり（本実施形態の場合、検索キーワード「肉野菜」による検索結果ページの閲覧者が多く）、影響力が大きいと判断できる。

【0203】

このため、本実施の形態において、顧客5は、「肉野菜」関連のコンテンツを増やすなどの対策を優先的に講じる必要があると判断でき、迅速かつ正確に閲覧者の希望に合うようにサイト内検索システム6a、WebページP₁～P_nを変更することができる。

【0204】

なお、本実施の形態において、ニーズ値の算出方法は、キーワード抽出機能14、ニーズ値算出機能21の設定を変更することにより、自由に変更できる。例えば、検索キーワードが入力されて得られたキーワード関連ページをアクセスした顧客サイト閲覧者8aの人数又はキーワード関連ページのアクセス数に基づいてニーズ値を算出してもよい。また、例えば、アクセスシーケンスにしたがったアクセスを行った閲覧者の人数、アクセスシーケンスにしたがったアクセスの実行された回数（閲覧回数）に基づいてニーズ値を算出してもよい。

【0205】

なお、量的評価値の例として、ニーズ値の他に機会値もある。機会値とは、Webサイトがその目的を遂行するに当たって、検索キーワードに関連する観点で、どれだけの事業機会を得たかを示す指標である。

【0206】

例えば、食品を販売しているECサイトにおいて、「砂糖」という検索キーワードによって閲覧してきた閲覧者が多数いれば、それだけ多くの砂糖を販売する事業機会を得たことになる。

【0207】

また、検索キーワードごとの入力回数のような、ニーズ値の算出に用いられた数量に基づいて機会値を算出してもよい。

【0208】

検索キーワードごとの入力回数をニーズ値とするか機会値とするかについては、Webサイトや閲覧者の特性、分析者の観点に依存する。

【0209】

例えば、商品そのものを販売するというよりは、商品に関する情報を提供し、その反応を知るために運営されているWebサイトにおいては、検索キーワードごとの入力回数をニーズ値として取り扱うことが適切である。

【0210】

一方、商品そのものをオンラインで販売することを主たる目的としているWebサイトでは検索キーワードごとの入力回数を機会値として取り扱い、各検索キーワードに関連した商品の販売の機会の量を示す指標とすることが適切である。

【0211】

（第3の実施の形態）

本実施の形態においては、上記第2の実施の形態に係る検索キーワード分析システムの変形例について説明する。

【0212】

図7は、本実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図である。

【0213】

本実施の形態に係る検索キーワード分析システム24は、記録媒体25に記録されている検索キーワード分析プログラム26を読み出して実行する。

【0214】

検索キーワード分析プログラム26は、システム24に読み込まれ、実行されると、ログ前処理機能12、アクセスシーケンス抽出機能13、キーワード抽出機能14、評価値算出機能39に加えて、グラフ化機能27を実現する。

【0215】

グラフ化機能27は、満足度・ニーズ値関連付け機能22によって生成された関係データを読み込む。

【0216】

そして、グラフ化機能27は、検索キーワードごとのニーズ値・満足度の値をグラフで表現した出力データ28を作成する。

【0217】

グラフ化機能27は、所定のグラフ化規則とグラフ化に利用されるパラメータとにしたがって、検索キーワードごとに満足度とニーズ値とが関連付けされた関係データをグラフ化する。

【0218】

例えば、グラフ化機能27は、散布図を作成する。この場合、満足度とニーズ値とをパラメータとし、散布図の一方の軸（X軸）をニーズ値、他方の軸（Y軸）を満足度とし、関係データにおけるニーズ値・満足度の値に基づいて各検索キーワードを表す点をプロットする。

【0219】

図8は、検索キーワード毎の満足度とニーズ値とを表すグラフの一例である。

【0220】

本実施の形態においては、検索キーワードごとの満足度とニーズ値とがグラフ化されて表示されるため、顧客5に視覚的な分析結果が提供される。

【0221】

したがって、顧客5は、より人間が直感的に解釈・判断しやすい分析結果を得ることができる。

【0222】

例えば、顧客5は、上記図8のグラフから、検索キーワード「りんご」についてはニーズ値が高く満足度も高いのに対して、検索キーワード「スーパー」「肉

野菜」についてはニーズ値が高いにもかかわらず満足度が低く、早急にサイト内検索システム6a、WebページP₁～P_nを変更し、対策を取る必要があると解釈される。

【0223】

なお、本実施の形態におけるグラフ化機能27の設定を変更することにより、グラフの軸の値を変換でき、グラフ表示効果を向上させることができる。例えば、グラフ化機能27は、ニーズ値、満足度を0から1の間に正規化するとしてもよい。また、ニーズ値、満足度に対数処理を施してグラフ表示してもよい。さらに、グラフ化機能27は、ニーズ値、満足度を棒グラフで表してもよい。さらに、グラフ化機能27は、グラフに適切な補助線を加えてもよい。

【0224】

(第4の実施の形態)

本実施の形態においては、上記第1乃至第3の実施の形態に係る検索キーワード分析システムの変形例について説明する。なお、以下においては、上記第1の実施の形態についての変形例を説明するが、他の第2、第3の実施の形態についても同様の変形が可能である。

【0225】

なお、本実施の形態においては、質的評価値のうち成功度を求める場合について説明するが、成功度ではない他の質的評価値を求める場合も同様である。

【0226】

図9は、本実施の形態に係る検索キーワード分析システムの一例を示すブロック図である。

【0227】

本実施の形態に係る検索キーワード分析システム29は、記録媒体30に記録されている検索キーワード分析プログラム31を読み出して実行する。

【0228】

検索キーワード分析プログラム31は、システム29に読み込まれ、実行されると、ログ前処理機能12、アクセスシーケンス抽出機能13、キーワード抽出機能14、評価値算出機能40を実現する。

【0229】

また、評価値算出機能40は、成功度算出機能32、成功度・キーワード関連付け機能41を含む。

【0230】

成功度算出機能32は、質的評価規則にしたがって、アクセスシーケンス抽出機能13によって算出されたアクセスシーケンスデータとキーワード抽出機能14によって算出されたキーワードデータとに基づいて、成功度を算出する。

【0231】

本実施の形態に係る質的評価規則として、例えば以下のようない内容の規則が定義される。

【0232】

顧客5は、代表的な検索キーワード（分析候補の検索キーワード）に対して、閲覧者にアクセスして欲しいWebページ（以下、「アクセス希望ページ」という）の識別情報を関連付けたリスト（以下、「アクセス希望ページリスト」という）を作成する。

【0233】

そして、顧客サイト閲覧者8aが検索結果ページ上のリンクにしたがってアクセス希望ページをアクセスしたら成功度をカウントアップする。

【0234】

ただし、検索結果ページからアクセス希望ページまでのリンククリック数が増加するにつれて成功度を減少させる。

【0235】

なお、顧客5が指定しなかった検索キーワードは「（その他）」とし、成功度は一切加算せず評価対象外とする。

【0236】

以上述べた質的評価規則において利用されるアクセス希望ページリストの一例を、表14に示す。

【0237】

【表14】

表14 アクセス希望ページリスト

検索キーワード	アクセス希望ページ
りんご	/apple/index.html
apple	/apple/index.html
肉野菜	/meat/index.html
肉野菜	/vegetable/index.html
...	...
(その他)	

【0238】

なお、一つの検索キーワードに対して複数のアクセス希望ページが対応付けされていてもよく、複数の検索キーワードが一つのアクセス希望ページに対応付けされていてもよい。

【0239】

また、表15に、検索結果ページからアクセス希望ページまでのリンククリック数が増加するにつれて成功度を減少させるための成功度重み付け表の一例を示す。

【0240】

【表15】

表15 成功度重み付け規則

検索結果ページ からのクリック数	成功度
1	100
2	80
3	60
4	40
5 以上	20
到達せず	0

【0241】

成功度算出機能32は、アクセスシーケンス抽出機能13によって生成されたアクセスシーケンスデータと、キーワード抽出機能14によって生成されたキーワードデータと、アクセス希望ページリストと、成功度重み付け表とを読み込む

。

【0242】

なお、成功度算出機能32は、キーワードデータの全てを読み込むのではなく、キーワードデータのうち検索キーワードとアクセス識別子の組み合わせのみを読み込むとしてもよい。

【0243】

次に、成功度算出機能32は、キーワードデータから、アクセス希望ページリストの検索キーワードに含まれている検索キーワードとそのアクセス識別子(ViewID)を取り出す。

【0244】

次に、成功度算出機能32は、アクセスシーケンスデータにおいて、上記作業によって取り出されたアクセス識別子と関連付けされているアクセスシーケンス識別子を取り出す。

【0245】

次に、成功度算出機能32は、取り出されたアクセスシーケンス識別子に関するアクセス識別子とリクエストとをアクセスシーケンスデータから取り出し、リクエストに含まれている検索キーワードを取り出す。

【0246】

表16に、この段階における第1の中間生成データを示す。

【0247】

【表16】

表16 中間生成データ（1）

ViewID	VisitID	リクエスト	検索キーワード
View20	Visit6	/index.html	
View21	Visit6	/search.cgi	
View22	Visit6	/search.cgi?q=%E3%82%8A%E3%82%93%E3%81%94	りんご
View26	Visit6	/general/index.html	
View28	Visit6	/apple/index.html	
View37	Visit6	/apple/tsugaru.html	
View45	Visit7	/index.html	
View47	Visit7	/search.cgi	
View48	Visit7	/search.cgi?q=apple	apple
...

【0248】

次に、成功度算出機能32は、各検索キーワードに対するアクセス希望ページがその検索キーワードが入力されたアクセスシーケンスにおいてアクセスされているかを調査する。

【0249】

アクセス希望ページがアクセスされている場合、成功度算出機能32は、検索キーワードの入力から何クリック後にアクセス希望ページがアクセスされたかを調査する。

【0250】

なお、検索キーワードの入力前にアクセス希望ページがアクセスされていても、そのアクセスは成功度の算出において考慮しない。

【0251】

上記表16の例において、アクセス識別子「View22」に関連する検索キーワード「りんご」に対するアクセス希望ページ識別子は、「/apple/index.html」である。このアクセス希望ページ識別子「/apple/index.html」は、アクセス識別子「View22」と同一のアクセスシーケンス識別子「Visit6」と関連するアクセス識別子「View28」に関連付けされている。

【0252】

そして、アクセス希望ページ識別子「/apple/index.html」は、検索キーワード「りんご」が入力された2クリック後に、入力されている。

【0253】

次に、成功度算出機能32は、成功度と検索結果ページからのクリック数との関係に基づいて、成功度を算出する。

【0254】

上記の例及び上記表15を適用した場合、アクセスシーケンス識別子「Visit6」で表されるアクセスシーケンスでは、検索キーワード「りんご」に対して満足度「80」が与えられる。

【0255】

表17に、この段階における第2の中間生成データを示す。

【0256】

【表17】

表17 中間生成データ（2）

VisitID	検索キーワード	成功度
Visit6	りんご	80
Visit7	apple	100
Visit21	肉野菜	0
Visit23	スーパー	60
Visit30	りんご	20

【0257】

そして、成功度算出機能32は、検索キーワードごとに、成功度の平均値を求めた成功度算出データを作成する。

【0258】

成功度・キーワード関連付け機能41は、成功度算出機能32で生成された成功度算出データと、キーワード抽出機能14で生成されたキーワードデータを統合し、検索キーワードごとの成功度を算出し、成功度と検索キーワードとを関係付け、関係データを作成する。

【0259】

関係データは、例えば、項目「満足度」が項目「成功度」に変化する点を除い

て上記表11と同様の形式となる。

【0260】

以上説明した本実施の形態における成功度算出方法を用いることで、ある検索キーワードに対し顧客にアクセスしてもらいたいとと考えているWebページへ適切に誘導がなされているかを知ることができる。

【0261】

(第5の実施の形態)

本実施の形態においては、上記第1から第4の実施の形態に係る検索キーワード分析システムの変形例について説明する。

【0262】

本実施の形態においては、検索システムが顧客サイト6外にある場合の検索キーワード分析システムについて説明する。

【0263】

なお、以下においては、上記第1の実施の形態についての変形例を説明するが、他の第2から第4の実施の形態についても同様の変形が可能である。

【0264】

図10は、本実施の形態に係る顧客サイトの構成の一例を示す図である。

【0265】

顧客サイト61は一以上のWebページP₁～P_nを持つ。また、顧客サイト61におけるいくつかのWebページは、顧客5以外の第3の運営主体によって運営されている複数のサイト外の検索システム（以下、「サイト外検索システム」という）341～34nの検索対象となっている。

【0266】

サイト外検索システム341～34nは、検索キーワードを入力すると、キーワード関連ページへのリンクを含む検索結果ページを表示する。

【0267】

本実施の形態においては、顧客サイト61と検索システム341～34nの運営主体が異なるため、上記第1から第4の実施形態の場合と異なり検索システム341～34nのWebアクセスログを取得することは困難である。

【0268】

したがって、本実施の形態においては、検索システム341～34nのWebアクセスログではなく、顧客サイト61のWebアクセスログに含まれるレファラ情報をを利用して検索キーワード分析を行う。

【0269】

表18に、顧客サイト61のWebアクセスログの一例を示す。

【0270】

【表18】

表18 レファラ付Webアクセスログ

ホスト名	User ID	日付	時刻	リクエスト	終了コード	転送バイト数	レファラ情報	ユーザエージェント
pc001.000.co.jp	—	2002/2/14	18:56:14	/index.html	200	400	http://www.000.O.ne.jp/□□□...	MOOiiia/4.0 (compatible:MOIE...
ws20.△△△.co.jp	—	2002/2/14	18:57:30	/cgi/login.cgi	200	700	—	MOOiiia/4.7 [△△5LLL](X11...
pc001.000.co.jp	—	2002/2/14	18:58:01	/products/index.html	200	2400	http://www.×××.com/△△△...	WΔm/0.3+m□e-p21-26+mOe-1.5.2
pc001.000.co.jp	—	2002/2/14	18:58:40	/products/info/p001.pdf	200	3000	http://www.000.O.ne.jp/○○○○...	MOOiiia/4.7 [△△5LLL](X11...
ws20.△△△.co.jp	—	2002/2/14	18:59:10	/cgi/order.cgi	200	700	http://www.000.O.ne.jp/□□□...	MOOiiia/4.0 (compatible:MOIE...

【0271】

図11は、サイト外検索システム341から顧客サイト61への閲覧遷移の一例を示すブロック図である。

【0272】

閲覧者によって検索キーワード入力画面35に検索キーワードが入力されると、検索結果ページ36が作成される。この検索結果ページ36を示すURLには、検索キーワード入力画面35に入力された検索キーワードが含まれる場合がある。

【0273】

閲覧者によって検索結果ページ36のリンクがクリックされ、顧客サイト61のキーワード関連ページが閲覧された場合、検索キーワードを含む検索結果ページのURLであるレファラ情報とアクセスされた顧客サイト61のキーワード関連ページのURLであるリクエストとが関連付けされて顧客サイト61のWebアクセスログに記録される。

【0274】

本実施の形態に係る検索キーワード分析システムは、上記図1に示す構成と同様の構成を持つため、ここでは図1を参照して本実施の形態の特徴的な部分について説明する。

【0275】

本実施の形態に係る顧客サイト61のWebアクセスログには、レファラ情報が含まれる。

【0276】

本実施の形態に係る検索キーワード分析システムにおいては、アクセスシーケンス抽出機能、キーワード抽出機能、満足度抽出機能の動作が、上記第1の実施の形態に係る検索キーワード分析システム9と異なる。

【0277】

本実施の形態に係るアクセスシーケンス抽出機能は、アクセスシーケンス識別子の付与、アクセスシーケンスの区切りの判定などについては、上記第1の実施の形態に係るアクセスシーケンス抽出機能13と同様の処理を実行する。

【0278】

また、本実施の形態に係るアクセスシーケンス抽出機能は、各アクセスについてのレファラ情報に基づいて、レファラ情報がサイト外検索システムを示すか否か判定する。

【0279】

表19に、サイト外検索システムを表すURLパターンと検索キーワードを抽出するためのクエリーパターンとを関係付けたデータを示す。

【0280】

【表19】

表19 パターン

検索システムのURL	クエリーパターン
http://www.△△△.com/	keyword
http://www.□□□.ne.jp/	kw
http://www.○○○.co.jp/	query
http://www.×××.co.jp/	word
http://www.◎◎◎.com/	q

【0281】

本実施の形態に係るアクセスシーケンス抽出機能は、サイト外検索システムからのアクセスか否かを判定する場合に、各アクセスのレファラ情報が検索システムのURLパターンを含むか調査する。

【0282】

そして、本実施の形態に係るアクセスシーケンス抽出機能は、レファラ情報がサイト外検索システムを示すアクセスに対してレファラフラグを付与したアクセスシーケンスデータを作成する。

【0283】

表20に、本実施の形態に係るアクセスシーケンスデータの一例を示す。

【0284】

【表20】

表20 アクセスシーケンスデータ

ViewID	VisitID	VisitorID	日付	時刻	リクエスト	レファラ フラグ
View1	Visit1	Visitor1	2002/2/14	18:56:14	/AAA.html	1
View2	Visit2	Visitor2	2002/2/14	18:57:30	/BBB.cgi	0
View3	Visit1	Visitor1	2002/2/14	18:58:01	/CCC.html	0
View4	Visit2	Visitor2	2002/2/14	18:58:40	/GGG/DDD.cgi	0
View5	Visit3	Visitor1	2002/2/14	19:42:10	/EEE.html	1

【0285】

表20において、レファラフラグ「0」は、サイト外検索システムからのアクセスでない旨を示す。

【0286】

一方、レファラフラグ「1」は、サイト外検索システムからのアクセスである旨を示す。

【0287】

このレファラフラグを記録することにより、アクセスシーケンスの途中で検索が行われた場合であっても、検索後のアクセス状況を調査し、満足度を算出できる。

【0288】

本実施の形態に係るキーワード抽出機能は、ログ前処理機能12によって生成された前処理データに基づいて、レファラ情報が検索システムのURLパターンを含むアクセスを求める。

【0289】

そして、本実施の形態に係るキーワード抽出機能は、URLパターンが含まれているレファラ情報の中から、その検索システムのURLパターンに対応付けられたクエリーパターンに基づいて、閲覧者がサイト外検索システムにおいて入力した検索キーワードを抽出する。

【0290】

表21に、本実施の形態に係るキーワード抽出機能によって生成されたキーワ

ードデータの一例を示す。

【0291】

【表21】

表21 キーワードデータ

ViewID	レファラ情報に含まれる キーワード
View6	りんご
View21	apple
View43	banana
View56	砂糖
View81	カレイ

【0292】

本実施の形態に係る満足度算出機能は、アクセスシーケンス抽出機能によって生成されたアクセスシーケンスデータから満足度を算出する。

【0293】

満足度の算出方法としては、例えば上記第1の実施の形態で説明した規則を適用する。

【0294】

すなわち、本実施の形態に係る満足度算出機能は、本実施の形態に係るアクセスシーケンス抽出機能によって生成されたアクセスシーケンスデータを読み込んだ後、Webページのアクセスごとに検索後閲覧時間を計算する。

【0295】

検索後閲覧時間は、同一のアクセスシーケンス識別子で表されるアクセスの中で、検索システムを利用したレファラ情報を持つアクセス（上記表20においてフラグ「1」のアクセス）から最後に閲覧されたアクセスまでの時刻の差分を求めることで計算できる。

【0296】

なお、検索システムをレファラ情報として持つアクセスがないアクセスシーケンスに関しては検索後閲覧時間を記録しない。

【0297】

また、本実施の形態に係る満足度算出機能は、上記表8のような検索後閲覧時間と満足度とを関連付けたデータを読み込み、アクセスごとの満足度を算出する。

【0298】

ただし、検索後閲覧時間の記録がない場合は満足度を記録しない。

【0299】

そして、本実施の形態に係る満足度算出機能は、アクセス識別子毎にアクセスシーケンス識別子と満足度とを関連付け、上記表9と同様な満足度算出データを作成する。

【0300】

以上説明した本実施の形態においては、顧客サイト61が検索システムを具備していない場合であっても、検索キーワードの満足度を算出することができる。

【0301】

したがって、検索キーワードの分析結果を得ることができる顧客サイトを大幅に増加することができる。

【0302】

なお、本実施の形態においては検索キーワード入力画面35がサイト外検索システム341にある場合について説明を行ったが、顧客サイト61内に検索キーワード入力画面があるとしてもよい。顧客が、検索業務をASP(Application Service Provider)サービスとして行っている業者に顧客サイト61内を検索するシステムを委託している場合などがこのようなケースに当てはまる。

【0303】

図12は、顧客サイト61内の検索キーワード入力画面96からサイト外検索システム341の検索結果ページ97へ移り、上記検索結果ページ97から顧客サイト61内のキーワード関連ページ98に遷移する一例を示すブロック図である。

【0304】

この図12において、閲覧者によって検索キーワード入力画面96に検索キーワードが入力されると、サイト外検索システム341により検索結果ページ97

が作成される。この検索結果ページ97を示すURLには、顧客サイト61内の検索キーワード入力画面96に入力された検索キーワードが含まれる場合がある。

【0305】

閲覧者によってサイト外検索システム341の検索結果ページ97のリンクがクリックされ、顧客サイト61のキーワード関連ページ98が閲覧された場合、検索キーワードを含む検索結果ページ97のURLであるレファラ情報とアクセスされた顧客サイト61のキーワード関連ページ98のURLであるリクエストとが関連付けされて顧客サイト61のWebアクセスログに記録される。

【0306】

(第6の実施の形態)

本実施の形態においては、第1の実施の形態に係る検索キーワード分析システムの変形例について説明する。

【0307】

第1から第5の実施の形態においては、アクセスシーケンスの抽出や質的評価値の算出と、検索キーワードの抽出を並列に実行し、最終的に関連付け機能によって統合する方法を用いていた。

【0308】

これに対し、本実施の形態では、検索キーワードの抽出処理と質的評価値の算出処理を直列に行う。

【0309】

なお、本実施の形態では、サイト内検索について分析を行う場合について記述しているが、本実施の形態と第5の実施形態とを組み合わせることでサイト外検索についての分析にも適用可能である。

【0310】

図13は、本実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図である。

【0311】

本実施の形態に係る検索キーワード分析システム42は、記録媒体43に記録

されている検索キーワード分析プログラム44を読み出して実行する。

【0312】

検索キーワード分析プログラム44は、システム42に読み込まれ、実行されると、ログ前処理機能12、アクセスシーケンス抽出機能13、キーワード抽出機能45、評価値算出機能46を実現する。

【0313】

キーワード抽出機能45、評価値算出機能46によって作成されたデータは、記録装置38に記録される。キーワード抽出機能45、評価値算出機能46は、記録装置38に記録されたデータを読み込み可能である。

【0314】

キーワード抽出機能45は、アクセスシーケンス抽出機能13によって生成されたデータから、顧客サイト閲覧者8aがサイト内検索システム6aにおいて入力した検索キーワードを抽出する。検索キーワードの抽出には上記第1の実施の形態と同等の方法を用いる。

【0315】

すなわち、キーワード抽出機能45は、前処理データから検索システムの場所に関する情報に基づいて、検索システムに対するアクセスを抽出する。

【0316】

次に、キーワード抽出機能45は、クエリーパターンに関する情報に基づいて、検索キーワードを抽出する。なお、エンコードされたパラメータについては判読可能な文字列へのデコードを行う。

【0317】

そして、キーワード抽出機能45は、アクセスシーケンス・キーワードデータを生成する。この際、検索システムを利用していないうアクセスに関してはキーワード項目に何も記録しない。

【0318】

表22に、アクセスシーケンス・キーワードデータの具体例を示す。

【0319】

【表22】

表22 アクセスシーケンス・キーワードデータ

ViewID	VisitID	VisitorID	日付	時刻	検索キーワード
View1	Visit1	Visitor1	2002/2/14	18:56:14	
View2	Visit2	Visitor2	2002/2/14	18:57:30	りんご
View3	Visit1	Visitor1	2002/2/14	18:58:01	
View4	Visit2	Visitor2	2002/2/14	18:58:40	
View5	Visit3	Visitor1	2002/2/14	19:42:10	たまねぎ

【0320】

評価値算出機能46は、アクセスシーケンス・キーワードデータに基づいて、質的評価値、例えば満足度を算出する。質的評価規則は第1の実施の形態と同様の規則を用いる。

【0321】

すなわち、評価値算出機能46は、アクセスシーケンス・キーワードデータを読み込み、アクセスシーケンスごとに検索後閲覧時間を計算する。検索後閲覧時間は、同一のアクセスシーケンス識別子によるアクセスの中で、キーワード関連ページへのアクセスと一番新しいアクセスの時刻の差分から計算される。

【0322】

次に、評価値算出機能46は、上記表8のような検索後閲覧時間と満足度の対応データを読み込み、アクセスシーケンスごとの満足度を算出する。

【0323】

表23に、この段階における中間生成データを示す。

【0324】

【表23】

表23 中間生成データ

ViewID	VisitID	満足度	検索キーワード
View1	Visit1	0	
View2	Visit2	20	りんご
View3	Visit1	0	
View4	Visit2	20	
View5	Visit3	10	たまねぎ

【0325】

次に、評価値算出機能46は、アクセスシーケンスについて満足度と検索キーワードを求める。この際、検索キーワードの記録のないアクセスシーケンスは処理対象としない。

【0326】

そして、評価値算出機能46は、検索キーワードごとに、満足度の平均を求め、関係データを生成する。

【0327】

関係データの具体的な内容は、上記表11と同様である。

【0328】

なお、本実施の形態において、評価値算出機能46は、量的評価値、例えばニーズ値を算出するとしてもよい。このニーズ値の算出方法は、上記第2の実施の形態で説明した方法が適用できる。また、評価値算出機能46によって質的評価値と量的評価値とを算出した場合、上記第3の実施の形態と同様にグラフ化機能27によって関係データに基づくグラフを表示するとしてもよい。

【0329】

以上説明した本実施の形態においては、アクセスシーケンス抽出、検索キーワード抽出、質的評価値算出及び量的評価値算出算出に関する各処理が並列ではなく直列に実行されており、上記各実施の形態と同様の効果を得ることができる。また、本実施の形態においては、ログを一括ではなく1アクセスシーケンスごとに処理することも可能になる。

【0330】

なお、上記各実施の形態においては、一つのアクセスシーケンスに対して検索キーワードが一回入力された場合を想定して説明しているが、一つのアクセスシーケンスに対して検索キーワードが複数回入力された場合であっても検索キーワードの質的評価値、量的評価値を求めることができる。

【0331】

例えば、一つのアクセスシーケンスに対して検索キーワードが複数回入力されている場合、そのアクセスシーケンスを検索キーワードの入力を境界として複数

のアクセスシーケンスを抽出し、この抽出されたアクセスシーケンスを用いて質的評価値、量的評価値を求めるとしてもよい。

【0332】

この場合、一つのアクセスシーケンスに対して検索キーワードが一回入力された場合における上記各実施の形態の処理方法を、細分化後的一つのアクセスシーケンスに対して適用することができる。

【0333】

(第7の実施の形態)

本実施の形態においては、分析対象のWebサイトのWebアクセスログ分析を実行し、分析対象ではない他のWebサイトに存在する外部コンテンツから分析対象のWebサイトが閲覧される傾向を分析するための検索キーワード分析システムについて説明する。

【0334】

本実施の形態では、外部コンテンツが検索システムであり、検索システムで入力された検索キーワードに基づいて閲覧の傾向を調査する。

【0335】

図14は、本実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図である。

【0336】

検索キーワード分析システム47は、記録媒体48に記録された検索キーワード分析プログラム49を読み込み、実行することにより、ログ前処理機能12a、アクセスシーケンス抽出機能13a、キーワード抽出機能14a、評価値算出機能39a、カテゴリ形成機能52、トレンド分析機能53、出力処理機能54を実現する。また、検索キーワード分析システム47は、出力装置55、記録装置38、58～60を具備する。

【0337】

なお、ログ前処理機能12a、アクセスシーケンス抽出機能13a、キーワード抽出機能14a、評価値算出機能39a、カテゴリ形成機能52、トレンド分析機能53、出力処理機能54は、ハードウェアにより実現されてもよい。

【0338】

本実施の形態において、例えば、記録装置38への記録、読み出しを経由して各機能12a～14a、39a、52～54の間でデータの提供、受け付けが行われるとしてもよい。

【0339】

記録装置38、58～60としては、例えば内部メモリ、ハードディスクなどが適用される。各記録装置38、58～60は、データベースを構成する要素であり、各データはデータベースに管理されるとしてもよい。また、各記録装置38、58～60は、自由に統合可能である。

【0340】

記録装置4に記録されているWebアクセスログ7は、閲覧者によってアクセスされた部分を示すアクセス先識別データ、アクセスされた時間を示す時間データ（例えば日付、時刻）、閲覧者が検索システムによる検索結果に基づいてアクセスした場合の検索キーワードとを関連付けたデータを含む。

【0341】

ログ前処理機能12a、アクセスシーケンス抽出機能13a、キーワード抽出機能14a、評価値算出機能39aは、上述したログ前処理機能12、アクセスシーケンス抽出機能13、キーワード抽出機能14、評価値算出機能39と同様の処理を実行するが、Webアクセスログ7の時間データに基づいて、所定の期間毎のデータに対して処理を実行し、所定の期間毎に、分析対象のWebサイトをアクセスするために入力された各検索キーワードと質的評価値である満足度と量的評価値であるニーズ値とを対応付けた関係データを作成する点が異なる。

【0342】

なお、上記各機能12a～14a、39aによって、所定の期間毎に、関係データを求める処理を実行するためには、例えばログ前処理機能12aが、Webアクセスログ7のうち所定の期間毎のデータを受け付けることで可能となる。

【0343】

評価値算出機能39aは、所定の期間毎に、分析対象のWebサイトをアクセスするために入力された各検索キーワードの満足度とニーズ値とを求めて、カテゴ

リ形成機能52とトレンド分析機能53に提供する。なお、例えば、同一期間内に、同一の検索キーワードの質的評価値又は量的評価値が複数算出された場合には、評価値算出機能39aは、その期間について、同一の検索キーワードの質的評価値又は量的評価値の平均を求めるとしてもよい。

【0344】

カテゴリ形成機能52は、記録装置58に記録されているカテゴリデータ72、所定の期間毎の各検索キーワードの満足度及びニーズ値に基づいて、各検索キーワードの属するカテゴリ単位で満足度及びニーズ値を求め、求めた各カテゴリの満足度及びニーズ値をトレンド分析機能53に提供する。

【0345】

カテゴリデータ72は、例えば、検索キーワードとこの検索キーワードの属するカテゴリとを関連付けたデータである。

【0346】

カテゴリ形成機能52は、例えば、期間毎に、カテゴリに属する検索キーワードのニーズ値の総和又は平均値を求め、期間毎のカテゴリのニーズ値とする。

【0347】

カテゴリ形成機能52は、例えば、期間毎に、カテゴリに属する各検索キーワードの満足度をニーズ値で重み付けした重み満足度の総和を、カテゴリに属する各検索キーワードのニーズ値の総和で割った値を求め、期間毎のカテゴリの満足度とする。

【0348】

トレンド分析機能53は、期間経過にそって、各検索キーワード及び各カテゴリのニーズ値と満足度の遷移状況を表すトレンドデータ73を求め、記録装置59に記録する。

【0349】

トレンド分析機能53は、領域計算機能53a、領域判断機能53b、メッセージ付加機能53cを具備する。

【0350】

領域計算機能53aは、記録装置60に記録されているニーズ境界値74、満

足度境界値 75、評価値算出機能 39a によって求められた期間毎の検索キーワードの満足度及びニーズ値、カテゴリ形成機能 52 によって求められた期間毎のカテゴリの満足度及びニーズ値に基づいて、期間毎に、検索キーワード及びカテゴリがニーズ境界値 74、満足度境界値 75 によって決定されるどの領域に位置するか判断する。

【0351】

ニーズ境界値 74、満足度境界値 75 は、ニーズ値、満足度の意味が切り換わる境界を表す値である。このニーズ境界値 74、満足度境界値 75 によって、分析対象のWeb サイトをアクセスするために入力された検索キーワード及びカテゴリの位置する領域が決定される。

【0352】

領域判断機能 53b は、期間の切り換わりに応じて、すなわち期間の経過にそ

って検索キーワードが領域を移動したか、カテゴリが領域を移動したか判断する。

【0353】

メッセージ付加機能 53c は、記録装置 60 に記録されているメッセージデータ 76 に基づいて、期間経過にそって領域を移動した検索キーワード又はカテゴリに、その移動した領域に対応するメッセージ（例えばコメント文、注釈など）を付する。

【0354】

メッセージデータ 76 は、例えば分析対象のWeb サイトをアクセスするために入力された検索キーワードの移動前の領域と移動後の領域とメッセージとを関連付けたデータである。なお、メッセージデータ 76 には、領域が変化した場合に、期間経過後の領域に移動したことを意味するメッセージを付するための規則が定義されていればよい。

【0355】

トレンド分析機能 53 は、期間毎の検索キーワード及びカテゴリのニーズ値及び満足度、期間毎の検索キーワード及びカテゴリの位置する領域、領域移動の有無を示す判断結果、領域を移動した検索キーワードに付されたメッセージ、領域

を移動したカテゴリに付されたメッセージを含むトレンドデータ73を記録装置59に記録する。

【0356】

出力処理機能54は、記録装置59に記録されているトレンドデータ73を、出力装置55を用いて出力するための処理を実行する。出力処理機能54は、トレンドデータ73を例えばテーブル、グラフ、各種の図などの形式で出力する。

【0357】

図15は、本実施の形態に係る検索キーワード分析システム47の動作の一例を示すフローチャートである。

【0358】

ステップV1において、ログ前処理機能12a、アクセスシーケンス抽出機能13a、キーワード抽出機能14a、評価値算出機能39aは、Webアクセスログ7の時間データに基づいて、所定の期間毎に、分析対象のWebサイトをアクセスするために入力された各検索キーワードと満足度とニーズ値とを対応付けた関係データを作成する。

【0359】

ステップV2において、カテゴリ形成機能52cは、期間毎に、各検索キーワードの属するカテゴリのニーズ値、満足度を求める。

【0360】

ステップV3において、トレンド分析機能53の領域計算機能53aは、ニーズ境界値74、満足度境界値75に基づいて、期間毎に、検索キーワード及びカテゴリの位置する領域を求め、検索キーワード及びカテゴリのニーズ値及び満足度、検索キーワード及びカテゴリの位置する領域をトレンドデータ73に含める。

【0361】

ステップV4において、トレンド分析機能53の領域判断機能53bは、期間経過にそって、検索キーワードの位置する領域が変化したか、カテゴリの位置する領域が変化したか判断し、判断結果をトレンドデータ73に含める。

【0362】

検索キーワード又はカテゴリの位置する領域が変化したと判断された場合、ステップV 5において、トレンド分析機能5 3のメッセージ付加機能5 3 cは、メッセージデータ7 6に基づいて、領域が変化した検索キーワード、領域が変化したカテゴリに、領域の変換に応じたメッセージを付し、結果をトレンドデータ7 3に含める。

【0363】

ステップV 6において、出力処理機能5 4は、分析者の操作に応じて、トレンドデータ7 3に基づいて、検索キーワード、カテゴリのニーズ値、満足度の遷移状況とメッセージとを表示する。

【0364】

以下、評価値算出機能3 9 aによって求められる関係データについて説明する。

【0365】

評価値算出機能3 9 aは、期間毎に、図1 6に示すような、分析対象のWebサイトをアクセスするために入力された検索キーワードのニーズ値と満足度を計算した結果を示す関係データを、トレンド分析機能5 3の領域計算機能5 3 a及びカテゴリ形成機能5 2に提供する。

【0366】

カテゴリ形成機能5 2は、期間毎に、図1 7に示すように、カテゴリデータ7 2と検索キーワードのニーズ値及び満足度とに基づいて、検索キーワード単位のニーズ値、満足度からカテゴリ単位のニーズ値、満足度を求め、カテゴリのニーズ値、満足度をトレンド分析機能5 3の領域計算機能5 3 aに提供する。

【0367】

トレンド分析機能5 3は、図1 8に示すような検索キーワードの遷移状況にメッセージを付したデータ、及び図1 9に示すようなカテゴリの遷移状況にメッセージを付したデータをトレンドデータ7 3に含め、トレンドデータ7 3を記録装置5 9に記録する。

【0368】

以下に、カテゴリ形成機能5 2について詳細に説明する。

【0369】

図20は、カテゴリ形成機能52の動作の一例を示すフローチャートである。

【0370】

ステップW1において、カテゴリ形成機能52は、カテゴリデータ72に基づいて、期間毎に、あるカテゴリに属する検索キーワードを識別する。

【0371】

ステップW2において、カテゴリ形成機能52は、期間毎に、(1)式に示すようにカテゴリに属する各検索キーワードのニーズ値の総和を求め、カテゴリのニーズ値とする。

【0372】**【数1】**

$$\text{カテゴリのニーズ値} = \sum (\text{カテゴリの } i \text{ 番目の検索キーワードのニーズ値}) \cdots (1)$$

【0373】

ステップW3において、カテゴリ形成機能52は、期間毎に、(2)式に示すように、カテゴリに属する各検索キーワードの満足度の加重平均を求め、カテゴリの満足度とする。

【0374】

【数2】

$$\begin{aligned}
 \text{カテゴリの満足度} &= \frac{\sum_{i=1}^n \{(\text{カテゴリの } i\text{ 番目の検索キーワードのニーズ値}) \times (\text{カテゴリの } i\text{ 番目の検索キーワードの満足度})\}}{\sum_{i=1}^n (\text{カテゴリの } i\text{ 番目の検索キーワードのニーズ値})}
 \end{aligned}$$

… (2)

【0375】

この数式 (2) では、カテゴリの満足度は、そのカテゴリに属する各検索キー

ワードの満足度とニーズ値との積の総和を、そのカテゴリに属する検索キーワードのニーズ値の総和で割って求められる。

【0376】

ステップW4において、カテゴリ形成機能52は、期間毎に、全てのカテゴリについてニーズ値及び満足度を求めたか判断し、全てのカテゴリについてニーズ値及び満足度を求めていない場合、上記ステップW1以下の処理を繰り返す。

【0377】

以下に、カテゴリの満足度を、そのカテゴリに属する各検索キーワードの満足度をそれぞれのニーズ値で重み付けした値の総和を、そのカテゴリに属する各検索キーワードのニーズ値の総和で割って求める理由について説明する。

【0378】

例えば、あるカテゴリCにおいて、検索キーワードC₁のニーズ値は極めて小さいが、検索キーワードC₁の満足度は極めて大きいとする。この場合、検索キーワードC₁を用いて分析対象のWebサイトにアクセスした閲覧者数は極めて少ないが、この小人数の閲覧者だけは満足していると判断される。

【0379】

一方、同一カテゴリCにおいて、検索キーワードC₂のニーズ値は極めて大きいが、検索キーワードC₂の満足度は極めて小さいとする。この場合、検索キーワードC₂を用いて分析対象のWebサイトにアクセスした閲覧者数は極めて多いが、この多人数の閲覧者は満足していないと判断される。

【0380】

このような場合に、単にカテゴリCの満足度を、このカテゴリCに属する検索キーワードC₁、C₂の満足度の平均とすると、閲覧者数の少なかった検索キーワードC₁の満足度と閲覧者の多かった検索キーワードC₂の満足度とが同レベルに扱われる。したがって、満足していない閲覧者が大多数であるにもかかわらず、カテゴリは中程度の満足度と判断される。

【0381】

しかしながら、カテゴリCの満足度は、閲覧者数の少なかった検索キーワードC₁の満足度よりも閲覧者の数が多い検索キーワードC₂の満足度から影響を受け

るようすべきである。

【0382】

そこで、カテゴリCに属する各検索キーワードC₁、C₂の満足度をニーズ値で重み付けし、合計を求める。

【0383】

そして、このカテゴリCと他のカテゴリとの間で、それぞれの閲覧者数に関係なく等しい基準で満足度を評価するために、カテゴリCについて、重み付けした満足度の合計を、そのカテゴリCの全ての閲覧者数すなわちそのカテゴリCに属する検索キーワードのニーズ値の合計で割る。

【0384】

図21は、カテゴリ形成機能52によるデータの変換結果の一例を示す図である。

【0385】

カテゴリ形成機能52は、検索キーワードとその検索キーワードの属するカテゴリとを関連付けたカテゴリデータ72と、期間毎の検索キーワードのニーズ値と満足度を示す関係データ77を受け付ける。

【0386】

そして、カテゴリ形成機能52は、期間毎に、カテゴリのニーズ値と満足度とを求めたカテゴリライズデータ78を求める。この図21の例では、5種類の検索キーワードが2種類のカテゴリに集約されている。各カテゴリについては、個々の検索キーワードに対する分析と同様の分析を行うことができる。

【0387】

以下に、トレンド分析機能53について詳細に説明する。

【0388】

図22は、トレンド分析機能53の動作の一例を示すフローチャートである。

【0389】

ステップX1において、トレンド分析機能53の領域計算機能53aは、期間毎の全ての検索キーワード及びカテゴリについてそれぞれのニーズ値及び満足度と、ニーズ境界値74及び満足度境界値75とに基づいて、期間毎の全ての検索

キーワード及びカテゴリの位置する領域を求め、検索キーワード及びカテゴリのニーズ値及び満足度、領域をトレンドデータ73に含める。

【0390】

ステップX2において、トレンド分析機能53の領域判断機能53bは、期間経過の前後で、検索キーワード及びカテゴリの位置する領域を比較し、検索キーワード及びカテゴリの領域が変化するか判断し、領域の変化する検索キーワード又はカテゴリを求め、結果をトレンドデータ73に含める。

【0391】

検索キーワード又はカテゴリの位置する領域が変化する場合、ステップX3において、トレンド分析機能53のメッセージ付加機能53cは、メッセージデータ76に基づいて、領域の変化する検索キーワード又はカテゴリに対して、期間経過前と後の領域に対応するメッセージを関連付け、トレンドデータ73に含める。

【0392】

図23は、ニーズ境界値74と満足度境界値75とによって決定される領域の名称の一例を示す図である。

【0393】

ニーズ値についてニーズ境界値74、満足度について満足度境界値75が設定されており、平面が4分割され、4つの領域に名称が付されている。

【0394】

この例では、ニーズ境界値74及び満足度境界値75をともに100としている。そして、ニーズ値と満足度の小さい領域を「失敗領域」、ニーズ値は小さいが満足度は大きい領域を「隠れ成功領域」、ニーズ値は大きいが満足度は小さい領域を「チャンスロス領域」、ニーズ値と満足度の大きい領域を「成功領域」としている。

【0395】

図24は、トレンド分析機能53によって作成されるトレンドデータ73の一例を示す図である。

【0396】

例えば、期間経過の切り換わり前と後で領域が変化した検索キーワード又はカテゴリには、移動先の領域に遷移したことを意味するメッセージが付される。この23のトレンドデータでは、領域変動のあった検索キーワード「食べ放題」「ランチ」にそれぞれメッセージ「成功領域へ遷移」「失敗領域へ遷移」というメッセージが付されている。

【0397】

このようなメッセージの付与は、期間経過後の領域とその領域に対応したメッセージとを関連付けたメッセージデータ76を用意し、トレンド分析機能53の領域判断機能53bによって期間経過前後で領域の変化した検索キーワード又はカテゴリを検出し、トレンド分析機能53のメッセージ付加機能53cによって、メッセージデータ76に基づいて、領域が変化した検索キーワード又はカテゴリに期間経過後に位置する領域に対応するメッセージを付することで可能である。

【0398】

図25は、領域が移動した場合に付与されるメッセージの他の例を示す図である。

【0399】

期間経過前後で境界を超えて変化する検索キーワード又はカテゴリに、どの境界をどの方向に超えたのか表すメッセージが付される。例えば、「隠れ成功領域」から「成功領域」に移動した検索キーワード又はカテゴリには、「ニーズ上昇」というメッセージを付する。他の場合も同様である。

【0400】

このようなメッセージの付与は、期間経過前の領域、期間経過後の領域、メッセージとを関連付けたメッセージデータ76を用意し、トレンド分析機能53の領域判断機能53bによって期間経過前後に領域の変化した検索キーワード又はカテゴリを検出し、トレンド分析機能53のメッセージ付加機能53cによって、メッセージデータ76に基づいて、領域の変化した検索キーワード又はカテゴリの期間経過前の領域と期間経過後に位置する領域に関連付けされたメッセージを求ることで可能である。

【0401】

以下に、出力処理機能54について詳細に説明する。

【0402】

出力処理機能54は、例えば図26に示すようなグラフを表示する。

【0403】

このグラフは、ニーズ値と満足度を表す座標軸を持ち、ニーズ境界値と満足度境界値とを表示し、各領域に領域名称を記述し、期間毎の検索キーワード又はカテゴリを示す指標をニーズ値と満足度に基づいてプロットし、プロットされた指標が線で結ばれて描かれている。なお、この図26では、期間経過毎の遷移状況を示す軌跡が矢印で表されている。

【0404】

その他にも、出力処理機能54は、例えば上記図18、19に示すようなニーズ値及び満足度に基づく期間毎の検索キーワード又はカテゴリのグラフ、上記図21に示すような期間毎のカテゴリズデータ78、上記図24に示すような検索キーワードの遷移状況とメッセージとを表すテーブル形式のトレンドデータ73を出力装置55に表示する。

【0405】

以上説明した本実施の形態においては、分析者は、分析対象のWebサイトについて集客戦略、コンテンツ改善の指針を得ることができる。

【0406】

例えば分析者は以下のよう効果を得ることができる。

【0407】

第1に、分析者は、閲覧者数は多いが閲覧者がすぐに分析対象のWebサイトの閲覧を終了するような検索キーワード及びカテゴリを調査でき、検索キーワード及びカテゴリに基づいて検索された分析対象のWebサイトが閲覧者の要求を満たすか否か判断できる。また、分析者は、閲覧者の要求を満たすことができないという一種のチャンスロスの状態を検出し、調査できる。したがって、分析者は、分析対象のWebサイトにおけるチャンスロスを改善し、分析対象のWebサイトの質を向上させることができる。

【0408】

第2に、分析者は、分析対象のWebサイトをアクセスする閲覧者の数を増加させ、かつアクセスした閲覧者を満足させる良好な検索キーワード及びカテゴリが存在するか判断できる。さらに、分析者は、このような検索キーワード及びカテゴリの良好な状態が一時的又は減少傾向にあるか判断でき、新たな検索キーワード及びカテゴリを検討すべきか判断できる。

【0409】

第3に、分析者は、個別の検索キーワード単位のみではなく、各検索キーワードの属するカテゴリ単位で静的又は動的な分析を行うことができる。このため、分析者は、自己の検討したい内容、検討したい事項に合わせて分析結果を得ることができ、分析対象のWebサイトについての閲覧傾向を適切かつ容易に把握でき、分析の精度を向上させることができる。

【0410】

第4に、分析者は、検索キーワード及びカテゴリの分析結果の時間的推移をグラフにより得ることができ、閲覧の傾向の時間的変化を容易に把握でき、分析作業を効率化できる。

【0411】

第5に、分析者は、検索キーワード及びカテゴリのトレンド分析の結果を用いて、膨大なデータから有用なデータを迅速に得ることができ、分析の精度を向上させることができ、分析に必要な労力を削減できる。

【0412】

第6に、分析者は、従来のWebアクセスログによる単純な集計ではなく、閲覧者の流行を容易かつ高精度に調査することができる。

【0413】

第7に、分析者は、検索キーワードを個別に分析しても把握できないようなカテゴリ全体のトレンドを把握できる。

【0414】

第8に、分析者は、閲覧者の意思で様々な形式で入力されつ検索キーワードについて容易かつ適切に自らの意思にそった分析結果を得ることができる。以下に

、この第8の効果について具体例を挙げて説明する。

【0415】

検索キーワードは、閲覧者の意思で様々な形式で入力されるため、Webアクセスログ分析により検索キーワードについて分析しても、分析者は自らの意思にそった分析結果を得ることが困難な場合がある。

【0416】

例えば、分析者は、料理関係の検索キーワード入力後に分析対象のWebサイトがどのようなアクセス状態か調べたいとする。この場合、従来のWebアクセスログ分析では、検索キーワード「料理」について閲覧者数を分析結果として求めることはできるが、料理と関係する他の検索キーワード「グルメ」「ファーストフード」などについての閲覧者数まで分析結果に含めることはできない。

【0417】

しかしながら、本実施の形態のようにカテゴリ単位で分析することによって、分析者は、自己の求める分析結果を適切に取得することができる。

【0418】

(第8の実施の形態)

本実施の形態においては、上記第7の実施の形態の変形例について説明する。本実施の形態に係る検索キーワード分析システムは、上記第7の実施の形態とトレンド分析機能が主に相違し、他の部分は同様であるとする。

【0419】

本実施の形態に係るトレンド分析機能は、ニーズ値境界値と満足度境界値とを用いることなく、ニーズ値と満足度を座標軸とする平面上における期間経過前後の検索キーワード及びカテゴリの移動距離と移動方向を求める。

【0420】

図27は、本実施の形態に係るトレンド分析機能の構成の一例を示すブロック図である。

【0421】

本実施の形態に係るトレンド分析機能79は、期間毎の検索キーワード及びカテゴリのニーズ値、満足度を座標値とし、期間経過前後における全ての検索キ

ワード及びカテゴリの移動距離を求め、移動距離の大きい検索キーワード及びカテゴリについて、その方向に応じたメッセージを付す。

【0422】

このトレンド分析機能79は、移動距離計算機能79a、移動距離判断機能79b、移動方向計算機能79c、メッセージ付加機能79dとを具備する。なお、トレンド分析機能79は、上記第7の実施の形態で説明したトレンド分析機能53と同じ機能も具備するとする。

【0423】

移動距離計算機能79aは、評価値算出機能39aによって求められた期間毎の検索キーワード、カテゴリ形成機能52によって求められた期間毎のカテゴリのニーズ値、満足度に基づいて、式（3）にしたがって、期間経過前後における全ての検索キーワード及びカテゴリの移動距離を求め、結果を移動距離判断機能79bに提供すると共にトレンドデータ81に含める。

【0424】

【数3】

… (3)

$$\text{移動距離} = \sqrt{\{(\text{後のニーズ値}) - (\text{前のニーズ値})\}^2 + \{(\text{後の満足度}) - (\text{前の満足度})\}^2}$$

【0425】

移動距離判断機能 79b は、記録装置 60 に記録されている移動距離しきい値 80 を超えて期間毎の検索キーワード又はカテゴリが移動したか判断すると共に

、結果をトレンドデータ81に含める。

【0426】

表24に、移動距離しきい値の具体例とトレンド分析機能79の動作の関係を示す。この表24の例では、移動距離しきい値が120と設定されており、移動距離が120以上の場合に、移動方向の計算及びメッセージ付与が行われることを示している。

【0427】

【表24】

(移動距離しきい値120の場合)

移動距離D	行なうべき動作
$D < 120$	—
$120 \leq D$	移動方向の計算、メッセージの付与

【0428】

移動方向計算機能79cは、移動距離が移動距離しきい値を超えた検索キーワード又はカテゴリのニーズ値、満足度に基づいて、式(4)にしたがって、期間経過前後の移動方向を計算し、結果をメッセージ付加機能79dに提供すると共にトレンドデータ81に含める。

【0429】

【数4】

…(4)

$$\text{移動方向} = \tan^{-1} \frac{(\text{後の満足度}) - (\text{前の満足度})}{(\text{後のニーズ値}) - (\text{前のニーズ値})}$$

【0430】

メッセージ付加機能79dは、記録装置60に記録されているメッセージデータ82に基づいて、移動距離しきい値を超えた検索キーワード又はカテゴリに対

して移動方向に応じたメッセージを付する。

【0431】

表25に、メッセージデータ82の一例を示す。移動方向を表す角度と、ニーズ及び満足の上昇、下降の意味するメッセージが対応付けられている。

【0432】

【表25】

移動方向 θ	付加するメッセージ
$0^\circ \leq \theta < 30^\circ$	ニーズ上昇
$30^\circ \leq \theta < 60^\circ$	ニーズ・満足度とも上昇
$60^\circ \leq \theta < 120^\circ$	満足度上昇
$120^\circ \leq \theta < 150^\circ$	ニーズ下降・満足度上昇
$150^\circ \leq \theta < 210^\circ$	ニーズ下降
$210^\circ \leq \theta < 240^\circ$	ニーズ・満足度とも下降
$240^\circ \leq \theta < 300^\circ$	満足度下降
$300^\circ \leq \theta < 330^\circ$	ニーズ上昇・満足度下降
$330^\circ \leq \theta < 360^\circ$	ニーズ上昇

【0433】

図28は、メッセージと移動方向の関係の具体例を示す。この図28において、移動方向 0° 以上～ 30° 未満の場合にはメッセージ「ニーズ上昇」と表示すると定義され、移動方向 30° 以上～ 60° 未満の場合にはメッセージ「ニーズ・満足度とも上昇」と表示すると定義され、移動方向 60° 以上～ 90° 未満の場合には「満足度上昇」と表示すると定義されている。

【0434】

なお、ニーズ値と満足度の双方が増加する移動方向 0° 以上～ 90° 未満の範囲のみではなく、ニーズ値と満足度の少なくとも一方が減少する 90° 以上～ 360° 未満範囲のメッセージも同様に定義できる。

【0435】

図29は、トレンド分析機能79によって作成されたトレンドデータ81の一例を示す図である。

【0436】

移動距離の大きい検索キーワード「食べ放題」「ランチ」に、移動方向に対応するメッセージが付されている。

【0437】

図30は、本実施の形態に係る検索キーワード分析システムの動作の一例を示すフローチャートである。

【0438】

ステップY1、Y2は、上記図15のステップV1、V2と同様であるため、説明を省略する。

【0439】

ステップY3において、トレンド分析機能79の移動距離計算機能79aは、期間経過前後における検索キーワード及びカテゴリの移動距離を求め、トレンドデータ81に含めるとともに移動距離判断機能79bに提供する。

【0440】

ステップY4において、トレンド分析機能79の移動距離判断機能79bは、記録装置60に記録されている移動距離しきい値80に基づいて、期間経過前後で検索キーワード及びカテゴリが移動距離しきい値80を超えるか判断し、結果をトレンドデータ81に含めるとともに移動方向計算機能79cに提供する。

【0441】

ステップY5において、トレンド分析機能79の移動方向計算機能79cは、移動距離が移動距離しきい値を超えたと判断された検索キーワード又はカテゴリの移動方向を計算し、結果をトレンドデータ81に含めるとともにメッセージ付加機能79dに提供する。

【0442】

ステップY6において、トレンド分析機能79のメッセージ付加機能79dは、記録装置60に記録されているメッセージデータ82に基づいて、移動方向の計算された検索キーワード又はカテゴリについてその移動方向に対応するメッセージを付加し、結果をトレンドデータ81に含める。

【0443】

ステップY7において、出力処理機能54は、分析者の操作に応じて、トレ

トデータ81に基づいて、検索キーワード又はカテゴリのニーズ値及び満足度の遷移状況とメッセージとを表示する。

【0444】

図31は、本実施の形態に係る出力処理機能54によって出力されたトレンドデータ81の一例を示す図である。

【0445】

この図31では、3期間以上でトレンド分析を行った結果得られたトレンドデータ81が表示されている。出力処理機能54は、連続する2期間毎に、トレンド分析機能79によって作成されたトレンドデータ81を同一平面上に表示する。この図31では、上記第1の実施の形態で説明したトレンドデータ73と本実施の形態で説明したトレンドデータ81とが組み合わされている。

【0446】

すなわち、この図31に示すグラフは、ニーズ値と満足度を表す座標軸を持ち、ニーズ境界値と満足度境界値とを表示し、各領域に領域名称を記述し、複数の期間毎の検索キーワード「宴会」を示す指標をそれぞれのニーズ値と満足度に基づいてプロットし、プロットされた指標が線で結ばれて描かれている。なお、この図31では、期間経過毎の遷移状況を示す軌跡が矢印で表されている。そして、図31では、領域の境界を超えた場合、及び移動距離の大きい場合のいずれかが成立した変化に対してメッセージが表示されている。

【0447】

以上説明した本実施の形態では、期間経過毎に、分析対象のWebサイトをアクセスするために入力される検索キーワードを分析でき、その変化量を移動距離、変化内容を移動方向として分析者に提供することができる。

【0448】

なお、上記第7及び本実施の形態においては、分析対象のWebサイトをアクセスするために入力される検索キーワードについて分析を行う場合について説明している。しかしながら、分析対象のWebサイトをアクセスする前にアクセスしていた外部コンテンツの存在する参照元のWebサイト又はドメインについて分析を行う場合にも同様に適用できる。

【0449】

その他にも、分析対象のWebサイトがキャンペーンメールに基づいてアクセスされた場合にこのキャンペーンメールについて分析を行い、キャンペーンの種別を表すIDについて分析結果を求めてよい。

【0450】

これにより、分析者は、検索キーワード、参照元のWebサイト、キャンペーンについてニーズの程度及び満足の程度を分析し、例えば空間上に表して結果を把握できる。

【0451】

また、上記第7及び本実施の形態の発明において、期間毎の関係データを求めるまでの処理は、上述した他の処理と変更可能である。すなわち、期間毎の関係データを求めるまでの処理として、上記各実施の形態で説明した各種手法を所定期間毎に実行することにより、同様の効果が得られる。

【0452】

また、上記第7及び本実施の形態の発明においては、カテゴリのニーズ値及び満足度を上記式（1）、式（2）で求めるとしているが、他の定義により他の手法でカテゴリのニーズ値及び満足度を求めるとしてもよい。

【0453】

また、本実施の実施の形態においては、検索キーワード及びカテゴリの移動距離及び移動方向を上記式（3）、式（4）で求めるとしているが、他の定義により他の手法で検索キーワード及びカテゴリの移動距離及び移動方向を求めるとしてもよい。

【0454】

また、上記第7及び本実施の形態においては、分析の対象はWebサイト単位でなくともよく、Webページ単位、複数のWebページから構成される任意のWebページグループ単位であってもよい。

【0455】

また、本実施の形態においては、移動距離が移動距離しきい値を超えた検索キーワード又はカテゴリのみ移動方向を算出するとしているが、全ての検索キーワード

ード又はカテゴリについて移動距離と移動方向を算出するとしてもよく、他の条件を満たす検索キーワード又はカテゴリについて移動距離と移動方向を算出するとしてもよい。

【0456】

また、上記第7及び本実施の形態発明においては、分析対象の閲覧される傾向の分析視点を、閲覧者が分析対象のWebサイトに期待する度合いを表すニーズ値と、分析対象のWebサイトが閲覧者に対して与える満足の度合いを表す満足度としている。しかしながら、閲覧の傾向の分析視点は、自由に変更できる。例えば、分析対象のWebサイトの目的が達成される機会の数と、目的が達成された数を表す達成度とを分析視点としてもよい。例えば、注文受付を目的とするWebサイトにおいて閲覧者の数を目標達成の機会数とし、注文を受け付けた数を達成度とし、この機会数と達成度を分析視点としてもよい。

【0457】

(第9の実施の形態)

本実施の形態においては、上記第1から第8までの実施の形態に係る検索キーワード分析システムの変形例について説明する。

【0458】

上記第1から第8までの実施の形態においては、抽出された一つのアクセスシーケンスに対して入力された検索キーワードが一つの場合について説明している。

【0459】

この上記第1から第8までの実施の形態のように、抽出されたあるアクセスシーケンスに対して一つの検索キーワードが入力されている場合のアクセスシーケンス、検索結果ページ2、顧客サイト61、顧客サイト61内のアクセスシーケンス、検索後閲覧時間の関係について、以下の図32から図34を用いて説明する。

【0460】

図32は、顧客サイト61内に検索システムが具備される場合のアクセスシーケンスの例を示す図である。このようなケースは、第1から第4の実施の形態及

び第6から第8の実施の形態において処理できる。

【0461】

この例では、閲覧者はWebページPa1、Pa2、検索結果ページSpa、WebページPa3～Pa5と順に閲覧を行い最後にWebページPa6を閲覧している。なお、WebページPa2、検索結果ページSpa、WebページPa3～Pa5は、顧客サイト61に属するページである。また、WebページPa2は検索キーワードが入力可能な画面である。

【0462】

この際、顧客サイト61内のアクセスシーケンスは、Pa2→Spa→Pa3→Pa4→Pa5となる。また、検索後閲覧時間は、上記第1の実施の形態の定義により、満足度を求めるアクセスシーケンス（すなわち顧客サイト61内のアクセスシーケンス）について最後のアクセス時刻から最初のキーワード関連ページ（すなわち検索結果ページSpaの次のページ）へのアクセス時刻を引いて得られる時間間隔であるため、WebページPa3からWebページPa5までのアクセス時間の間隔となる。

【0463】

図33は、検索システムが顧客サイト61外に存在する場合のアクセスシーケンスの例を示す図である。このようなケースは、上記第5の実施の形態において処理できる。

【0464】

この例では、閲覧者は、WebページPb1、検索結果ページSpb、WebページPb2～Pb6と順に閲覧を行い、最後にWebページPb7を閲覧している。なお、WebページPb2～Pb6は、顧客サイト61に属するWebページである。また、サイト外検索システム341において、閲覧者によって検索キーワードを入力可能な画面Pb1に検索キーワードが入力されると、検索結果ページSpbが作成される。検索結果ページSpbには、顧客サイト61のWebページPb2へのリンクが含まれている。

【0465】

この際、顧客サイト61内のアクセスシーケンスはPb2→Pb3→Pb4→

Pb5→Pb6となる。また、検索後閲覧時間は、上記第5の実施の形態の定義により、同一のアクセスシーケンス識別子で表されるアクセスの中で検索システムを利用したレファラ情報を持つアクセスから最後に閲覧されたアクセスまでの時刻の差分であるため、WebページPb2からWebページPb6までのアクセス時間の間隔となる。

【0466】

図34は、この図33と同様に、検索システムが顧客サイト61外にあり、さらに顧客サイト61内のアクセスシーケンスの途中に検索結果ページが存在する場合のアクセスシーケンスの例を示す図である。このようなケースは、上記第5の実施の形態において処理できる。

【0467】

この例では、閲覧者はWebページPc1～Pc3、検索結果ページSPc、WebページPc4～Pc6と順にアクセスを行い、最後にWebページPc7を閲覧している。なお、WebページPc2～Pc6は、顧客サイト61に属するWebページである。また、WebページPc3は、検索キーワードを入力可能な画面であり、その入力の結果に応じてサイト外検索システム341は、検索結果ページSPcを出力する。

【0468】

この際、顧客サイト61内のアクセスシーケンスはPc2→Pc3→Pc4→Pc5→Pc6となる。また、検索後閲覧時間は、上記第5の実施の形態の定義により、同一のアクセスシーケンス識別子で表されるアクセスの中で検索システムを利用したレファラ情報を持つアクセスから最後に閲覧されたアクセスまでの時刻の差分であるため、WebページPc4からWebページPc6までのアクセス時間の間隔となる。

【0469】

上記第1から第3、第5から第8の実施の形態では、上記図32から図34までに示すように、抽出されたアクセスシーケンスにおいて検索キーワードの入力が一回であり、検索結果ページから後のWebページの閲覧時間である検索後閲覧時間が満足度として処理が行われる。

【0470】

上記第4の実施の形態では、抽出されたアクセスシーケンスにおいて検索キーワードの入力が一回であり、この検索キーワードについての検索結果に対する成功度が求められる。

【0471】

本実施の形態においては、このような上記第1から第8の実施の形態をさらに発展させ、抽出されたあるアクセスシーケンスにおいて、検索キーワードが複数回入力されている場合について説明する。

【0472】

本実施の形態においては、顧客サイト内のアクセスシーケンスをさらに1つ以上のアクセスシーケンス（以下、「サブアクセスシーケンス」という）に細分化したものについて説明する。なお、状況によっては顧客サイト内のアクセスシーケンスにおいて、サブアクセスシーケンスが存在しない場合もある。

【0473】

そして、上記各実施の形態と同様の処理方法をサブアクセスシーケンスに対して適用する。

【0474】

サブアクセスシーケンスは、検索結果ページへのアクセスに基づいて定められる始端を抽出し、この抽出された始端と同一のアクセス要求側による連続的なアクセスのうち、始端以降のアクセスである終端によって定められる。顧客サイト内のアクセスシーケンスからサブアクセスシーケンスを抽出するサブアクセスシーケンス抽出規則の一例を以下に説明する。

【0475】

サブアクセスシーケンスの始端は、検索システムが顧客サイト内にある場合、第1規則を用いて抽出し、顧客サイト外にある場合、第2規則を用いて抽出する。

【0476】

第1規則として、抽出されたアクセスシーケンスにおいて顧客サイト内の検索システムが検索結果ページを出力している場合、サブアクセスシーケンスの始端

は、この顧客サイト内の検索システムによって出力された検索結果ページへのアクセスとする。

【0477】

第2規則として、抽出されたアクセスシーケンスにおいて顧客サイト外の検索システムが検索結果ページを出力し、しかも検索結果ページへアクセスの一つ後のアクセスが顧客サイト内のWebページへのアクセスである場合、サブアクセスシーケンスの始端は、顧客サイト外の検索システムが出力した検索結果ページをレファラ情報として持つ顧客サイト内のWebページへのアクセスとする。

【0478】

すなわち、顧客サイト外の検索システムによって検索が行われている場合、サブアクセスシーケンスの始端は、顧客サイト外の検索システムによる検索結果ページ上で選択された顧客サイト内のWebページへのアクセスとする。

【0479】

なお、「あるWebページAへのアクセスの一つ後のアクセス」とは、同一のアクセス要求側による連続的なアクセスのうち、あるWebページAへのアクセス以降に行われたアクセスで、しかもあるWebページAへのアクセスに時間的に最も近いアクセスを指し、例えば上記図33において、あるWebページを検索結果ページSPbとすると、WebページPb2へのアクセスが該当する。

【0480】

サブアクセスシーケンスの終端は、第3規則を用いて抽出する。

【0481】

第3規則として、サブアクセスシーケンスの終端は、サブアクセスシーケンスの始端からみて次のサブアクセスシーケンスの始端の一つ前の顧客サイト内のアクセスとする。ただし、サブアクセスシーケンスの始端からみて次のサブアクセスシーケンスの始端が存在しない場合は、顧客サイト内のアクセスシーケンスの終端を、そのサブアクセスシーケンスの終端とする。

【0482】

なお、「あるWebページAへのアクセスの一つ前のアクセス」とは、同一のアクセス要求側による連続的なアクセスのうち、あるWebページAへのアクセ

ス以前に行われたアクセスで、しかもあるWebページAへのアクセスに時間的に最も近いものを指す。また「あるWebページAへのアクセスの一つ前の顧客サイト内のアクセス」とは、同一のアクセス要求側による連続的なアクセスのうち、あるWebページAへのアクセス以前に行われたアクセスで、しかもあるWebページAへのアクセスに時間的に最も近い顧客サイト内のアクセスを指し、例えば上記図34において、あるWebページAをWebページPc4とするとWebページPc3が該当する。

【0483】

また、顧客サイト内の検索システムが出力した検索結果ページへのアクセスが、顧客サイト外の検索システムが出力した検索結果ページをレファラ情報を持つWebページへのアクセスにも該当する場合に第4規則を用いる。

【0484】

第4規則として、顧客サイト内の検索システムが出力した検索結果ページへのアクセスが、顧客サイト外の検索システムが出力した検索結果ページをレファラ情報を持つWebページへのアクセスにも該当する場合、顧客サイト内の検索システムが出力した検索結果ページへのアクセスをサブアクセスシーケンスの始端とする。

【0485】

あるアクセスシーケンスに対して複数の検索キーワードが入力されている場合のアクセスシーケンス、検索結果ページ、顧客サイト、顧客サイト内のアクセスシーケンス、サブアクセスシーケンスの関係について、以下の図35から図37を用いて説明する。

【0486】

図35は、顧客サイト61内に検索システムが具備されている場合のアクセスシーケンスにおいて複数の検索が行われた例を示す図である。この図35はサイト内検索システムにおいて検索が1回だけ行われた上記図32に対応付けられるものである。

【0487】

この例では、閲覧者はWebページPd1、Pd2、検索結果ページSPd、

WebページPd3～Pd5、検索結果ページSp e、Sp f、WebページPd6、Pd7と順に閲覧を行い、最後にWebページPd8を閲覧している。なお、WebページPd2～Pd7及び検索結果ページSp d～Sp fは、顧客サイト61に属するページである。また、WebページPd2、Pd5、及び各検索結果ページSp d～Sp fは検索キーワードを入力可能な画面である。

【0488】

この際、上記サブアクセスシーケンス抽出規則により、検索結果ページSp dへのアクセスから次の検索結果ページSp eへのアクセスの一つ手前の顧客サイト内のページへのアクセス(すなわちWebページPd5へのアクセス)までのSp d→Pd3→Pd4→Pd5がサブアクセスシーケンス100となる。

【0489】

同様に、検索結果ページSp eへのアクセスから次の検索結果ページSp fの一つ手前の顧客サイト61内のページへのアクセスまで(すなわち検索結果ページSp eへのアクセス)がサブアクセスシーケンス101となる。

【0490】

同様に、検索結果ページSp fへのアクセスから顧客サイト61内のアクセスシーケンスの終端のアクセス(すなわちWebページPd7へのアクセス)までのSp f→Pd6→Pd7がサブアクセスシーケンス102となる。

【0491】

なお、この例のように検索結果ページにおいて検索キーワードが入力可能な場合や、キーワード関連ページが検索結果ページである場合がある。このような場合、上記図45で示した画面遷移とは多少異なることとなる。

【0492】

図36は、検索システムが顧客サイト外に存在する場合のアクセスシーケンスにおいて複数の検索が行われた状態例を示す図である。この図36は、顧客サイト61外の検索システム341において検索が1回だけ行われた上記図34に対応付けられるものである。

【0493】

この例では、閲覧者はWebページPe1～Pe4、検索結果ページSp g、

WebページPe5～Pe7、検索結果ページSPh、WebページPe8～Pe10と順にアクセスを行い、最後にWebページPe11を閲覧している。なお、WebページPe2～Pe10は、顧客サイト61に属するWebページである。またWebページPe4、Pe7は、検索キーワード可能な画面であり、その入力の結果に応じてサイト外検索システム341は、検索結果ページSPg、SPhを出力する。

【0494】

この際、上記サブアクセスシーケンス抽出規則により、検索結果ページSPgをレファラとして持つページへのアクセス（すなわちWebページPe5へのアクセス）から次の検索結果ページSPhをレファラとして持つページへのアクセス（すなわちWebページPe8へのアクセス）の一つ手前の顧客サイト61内のページへのアクセス（すなわちWebページPe7へのアクセス）までのPe5→Pe6→Pe7が第1のサブアクセスシーケンスとなる。

【0495】

同様に、検索結果ページSPhをレファラとして持つページへのアクセス（すなわちWebページPe8へのアクセス）から顧客サイト61内のアクセスシーケンスの終端のアクセス（すなわちWebページPe10へのアクセス）までのPe8→Pe9→Pe10が第2のサブアクセスシーケンスとなる。

【0496】

図37は、顧客サイト内の検索システムと顧客サイト外の検索システムの双方を用いてアクセスシーケンスの途中で検索が行われた例を示す図である。この図37は、上記図32～34を組み合わせたものに対応付けられるものである。

【0497】

この例では、閲覧者はWebページPf1、検索結果ページSPi、WebページPf2、検索結果ページSPj、WebページPf3～Pf5、検索結果ページSPk、検索結果ページSPl、WebページPf6、検索結果ページSPm、SPn、WebページPf7、Pf8と順に閲覧を行い、最後にWebページPf9を閲覧している。なお、WebページPf2～Pf8及び検索結果ページSPj、SPl～SPnは、顧客サイト61に属するページである。また、W

WebページPf1、Pf2、Pf5、Pf6及び各検索結果ページSPi～SPnは、検索キーワードを入力可能な画面である。WebページPf1での検索キーワードの入力の結果に応じてサイト外検索システム341は検索結果ページSPiを出力し、WebページPf5での入力に応じてサイト外検索システム342は検索結果ページSPkを出力する。

【0498】

この際、上記サブアクセスシーケンス抽出規則により、検索結果ページSPiをレファラとして持つページへのアクセス（すなわちWebページPf2へのアクセス）から検索結果ページSPjへのアクセスの一つ手前の顧客サイト61内のページへのアクセスまで（すなわちWebページPf2のアクセス）がサブアクセスシーケンス103となる。

【0499】

同様に、検索結果ページSPjへのアクセスから検索結果ページSPkへのアクセスの一つ手前の顧客サイト61内のページへのアクセス（すなわちWebページPf5へのアクセス）までのSPj→Pf3→Pf4→Pf5がサブアクセスシーケンス104となる。このとき、顧客サイト61内の検索システムによる検索結果ページSP1へのアクセスが顧客サイト61外の検索システム342による検索結果ページSPkをレファラとして持つページへのアクセスにも該当しているため、上記サブアクセスシーケンス抽出規則により、顧客サイト61内の検索システムによる検索結果ページSP1へのアクセスをサブアクセスシーケンス105の始端としている。

【0500】

また同様に、検索結果ページSP1へのアクセスから検索結果ページSPmへのアクセスの一つ手前の顧客サイト61内のページへのアクセス（すなわちWebページPf6へのアクセス）までのSP1→Pf6がサブアクセスシーケンス105となる。

【0501】

以下同様に、検索結果ページSPmがサブアクセスシーケンス106、そしてSPn→Pf7→Pf8がサブアクセスシーケンス107となる。

【0502】

上記のようなサブアクセスシーケンス抽出規則に基づいて、顧客サイト61内のアクセスシーケンスから抽出されたサブアクセスシーケンスと検索後閲覧時間との関係について以下に説明する。

【0503】

顧客サイト61内のアクセスシーケンスにおける検索後閲覧時間は、検索システムが顧客サイト61内にある場合は「満足度を求めるアクセスシーケンス（すなわち顧客サイト61内のアクセスシーケンス）について最後のアクセス時刻から最初のキーワード関連ページ（すなわち検索結果ページの次のページ）へのアクセス時刻を引いて得られる時間間隔」であり、検索システムが顧客サイト61外にある場合は「検索システムを利用したレファラ情報を持つアクセスから最後に閲覧されたアクセスまでの時刻の差分」であった。これをサブアクセスシーケンスに当てはめてみると以下のようになる。

【0504】

図38は、顧客サイト61内の検索システムが検索結果ページを出力している場合において、検索結果ページへのアクセスが始端のサブアクセスシーケンスと検索後閲覧時間との関係の一例を示す図である。このサブアクセスシーケンスにおいて閲覧者は、検索結果ページS P o、W e bページP g 1～P g 3の順に閲覧している。このようなケースは、上記第1から第4の実施の形態及び第6から第8の実施の形態において処理できる。

【0505】

このような場合、検索結果ページS P oへのアクセスの次のアクセスであるW e bページP g 1へのアクセスからサブアクセスシーケンスの最後のアクセスであるW e bページP g 3へのアクセスまでの時間間隔を検索後閲覧時間とする。ただし、サブアクセスシーケンスが検索結果ページS P oからのみ構成されている場合、検索後閲覧時間は0とする。

【0506】

図39は、顧客サイト61外の検索システムが検索結果ページを出力する場合において、顧客サイト61外の検索システムによる検索結果ページをレファラに

持つページへのアクセスが始端のサブアクセスシーケンスと検索後閲覧時間との関係の一例を示す図である。このサブアクセスシーケンスにおいて閲覧者は、WebページPh1～Ph4の順に閲覧している。このようなケースは、上記第5の実施の形態において処理できる。

【0507】

このような場合、検索結果ページの次のアクセス（すなわち検索結果ページをレファラに持つページへのアクセス）であるWebページPh1へのアクセスからサブアクセスシーケンスの最後のアクセスであるWebページPh4へのアクセスまでの時間間隔を検索後閲覧時間とする。

【0508】

サブアクセスシーケンスの始端が顧客サイト61外の検索システムが出力した検索結果ページをレファラ情報に持つWebページへのアクセスである場合、サブアクセスシーケンスの始端のアクセスからサブアクセスシーケンスの終端までの閲覧時間を、検索後閲覧時間とする。ただし、サブアクセスシーケンスの始端とサブアクセスシーケンスの終端が同一の場合、検索後閲覧時間はサブアクセスシーケンスの始端のアクセス時間とする。

【0509】

本実施の形態において、顧客サイト61内の検索システムを用いて抽出されたサブアクセスシーケンスには、検索結果ページが含まれるが、顧客サイト61外の検索システムを用いて抽出されたサブアクセスシーケンスには、検索結果ページが含まれない。しかしながら、上記図38及び図39で示す関係に基づいて検索後閲覧時間を計算することにより、検索システムが顧客サイト61内であっても顧客サイト61外であっても、顧客サイト61内でのアクセスの状態についての質的評価値を公平に算出することができる。

【0510】

以下、本実施の形態の概念を、上記第1の実施の形態に適用した場合について説明する。しかしながら、本発明の概念は、上記第2から第8までのいずれの実施の形態に対しても同様に適用可能である。

【0511】

図40は、本実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図である。

【0512】

本実施の形態に係る検索キーワード分析システム83は、記録媒体84に記録されている検索キーワード分析プログラム85を読み出して実行する。

【0513】

検索キーワード分析プログラム85は、システム83に読み込まれ、実行されると、ログ前処理機能12、サブアクセスシーケンス抽出機能86、キーワード抽出機能87、評価値算出機能88を実現する。

【0514】

また、評価値算出機能88は、満足度算出機能89と満足度・キーワード関連付け機能16を含む。

【0515】

上記各機能12、86～88によって作成されたデータは、記録装置38に記録される。各機能12、86～88は記録装置38に記録されたデータを読み込み可能である。すなわち、各機能12、86～88間での各種データの受け渡しは、記録装置38を経由して行われてもよい。この場合、データの提供側は、提供するデータを記録装置38に記録し、データの受付側は、記録装置38からデータを読み出す。

【0516】

図41は、サブアクセスシーケンス抽出機能86の具体例を示すブロック図である。

【0517】

サブアクセスシーケンス抽出機能86は、アクセスシーケンス抽出機能90とサブアクセスシーケンス判定機能91を含む。

【0518】

アクセスシーケンス抽出機能90は、アクセスシーケンス抽出機能13と同等の機能を含むが、出力結果であるアクセスシーケンスデータにサブアクセスシーケンスを判定するために必要な情報を付加する特徴がある。

【0519】

アクセスシーケンス抽出機能90は、所定のアクセスシーケンス抽出規則にしたがって一連のページアクセスからアクセスシーケンスを取り出し、アクセスシーケンスの抽出結果を示すアクセスシーケンスデータを作成する。

【0520】

この際、例えばアクセスシーケンス抽出規則として上記第1の実施の形態で述べた規則を利用できる。

【0521】

表26に、アクセスシーケンス抽出機能90によって作成されたアクセスシーケンスデータの一例を示す。

【0522】

【表26】

表26 アクセスシーケンスデータ

ViewID	VisitID	VisitorID	日付	時刻	リクエスト	レファラ
View1	Visit1	Visitor1	2002/2/14	18:56:14	/AAA.html	http://www.□□□.ne.jp/search.cgi?kw=vegetable
View2	Visit2	Visitor2	2002/2/14	18:57:30	/BBB.cgi	—
View3	Visit1	Visitor1	2002/2/14	18:58:01	/CCC.html	http://www.○○○.co.jp/AAA.html
View4	Visit2	Visitor2	2002/2/14	18:58:40	/search.cgi?q=food	http://www.○○○.co.jp/BBB.html
View5	Visit1	Visitor1	2002/2/14	18:59:10	/EEE.html	—
...

【0523】

この表26に示すように、上記表3のアクセスシーケンスデータにレファラ情報を付加したデータが、本実施の形態に係るアクセスシーケンス抽出機能90によって作成されたアクセスシーケンスデータとなる。なお、この表26において、レファラ情報が存在しない場合は「-」で表している。

【0524】

サブアクセスシーケンス判定機能91は、抽出されたアクセスシーケンスからサブアクセスシーケンスを切り出す機能である。このサブアクセスシーケンス判定機能91は、上記表26に示すアクセスシーケンスデータに基づいてサブアクセスシーケンスを抽出する。

【0525】

図42は、サブアクセスシーケンス判定機能91の具体例を示すブロック図である。

【0526】

サブアクセスシーケンス判定機能91は、始端を判定するサブアクセスシーケンス始端判定機能92と、終端を判定するサブアクセスシーケンス終端判定機能93とを含む。

【0527】

サブアクセスシーケンス始端判定機能92は、上述したサブアクセスシーケンス抽出規則と上記表26に示すアクセスシーケンスデータとに基づき、サブアクセスシーケンスの始端を判定する。

【0528】

サブアクセスシーケンス始端判定機能92は、あるアクセスが顧客サイト61内の検索システムが出力した検索結果ページ2へのアクセスであることを判定するため、例えば上記表5のような顧客サイト61内の検索システムを特定するための情報94を用い、始端を求める。

【0529】

また、サブアクセスシーケンス始端判定機能92は、あるアクセスが顧客サイト61外の検索システムが出力した検索結果ページをレファラ情報を持つWebページへのアクセスであることを判定するために、例えば上記表19のような顧

客サイト61外の検索システムを特定するための情報95を用い、始端を求める

。

【0530】

次に、サブアクセスシーケンス終端判定機能93は、上述したサブアクセスシーケンス抽出規則、サブアクセスシーケンス始端判定機能92によって判定された始端データ、上記表26に示すアクセスシーケンスデータに基づいて、サブアクセスシーケンスの終端を判定する。

【0531】

そして、サブアクセスシーケンス終端判定機能93は、判定された始端と終端とにに基づくサブアクセスシーケンスの抽出結果を示すサブアクセスシーケンスデータを満足度算出機能88に提供する。

【0532】

具体的には以下のような動作をする。まず、サブアクセスシーケンス始端判定機能92は上記表26のアクセスシーケンスデータを同一のVisitIDを持つものごとにまとめ、日付・時刻の古い順に並び替える。表27に処理の結果生成された一時的な第1の表を示す。

【0533】

【表27】

表27 一時的な第1の表

ViewID	VisitID	VisitorID	Visitor	日付	時刻	リクエスト	レファラ
View1	Visit1	Visitor1	2002/2/14	18:56:14	/AAA.html	http://www.□□□.ne.jp/search.cgi?kw=vegetable	
View3	Visit1	Visitor1	2002/2/14	18:58:01	/CCC.html	http://www.○○○.co.jp/AAA.html	
View5	Visit1	Visitor1	2002/2/14	18:59:10	/EEE.html	http://www.○○○.co.jp/CCC.html	
View6	Visit1	Visitor1	2002/2/14	19:00:02	/search.cgi?q=food	http://www.○○○.co.jp/EEE.html	
View8	Visit1	Visitor1	2002/2/14	19:01:23	/FFF.html	http://www.○○○.co.jp/search.cgi?q=food	
View13	Visit1	Visitor1	2002/2/14	19:02:25	/GGG.html	http://www.○○○.co.jp/FFF.html	
View2	Visit2	Visitor2	2002/2/14	18:57:30	/BBB.cgi	-	
View4	Visit2	Visitor2	2002/2/14	18:58:40	/search.cgi?q=food	http://www.○○○.co.jp/BBB.html	
...	

【0534】

この例において、顧客サイト61のURLは、「<http://www.○○○.co.jp/>」から始まるとする。なお、この表27におけるアクセスシーケンスVisit1の画面遷移の状況を図32～37と同様の形式で表現したものを図43に示す。アクセスシーケンスVisit1では、顧客サイト61外の検索システムが出力した検索結果ページ「<http://www.□□□.ne.jp/search.cgi?kw=vegetable>」→顧客サイト61内のWebページ「AAA..html」→「CCC.html」→「EEE.html」→顧客サイト61内の検索システムが出力した検索結果ページ「<search.cgi?q=food>」→顧客サイト61内のWebページ「FFF.html」→「GGG.html」という順にアクセスが行われている。そして、上記サブアクセスシーケンス抽出規則により、「AAA..html」→「CCC.html」→「EEE.html」が第1のサブアクセスシーケンス108、「<search.cgi?q=food>」→「FFF.html」→「GGG.html」が第2のサブアクセスシーケンス109となる。以下具体的に、これらのサブアクセスシーケンスをどのように抽出するかについて説明を行う。

【0535】

上記並び替えが行われた後、上記サブアクセスシーケンス抽出規則をもとに、上記表27から始端となるアクセスを判定する。始端とされたものには始端フラグをつける。表28に処理の結果生成された一時的な第2の表を示す。

【0536】

【表28】

表28 一時的な第2の表

ViewID	VisitID	VisitorID	日付	時刻	リクエスト	レフアラ	始端 フラグ
View1	Visit1	Visitor1	2002/2/14	18:56:14	/AAA.html	http://www.□□□.ne.jp/search.cgi?KW=vegitable	1
View3	Visit1	Visitor1	2002/2/14	18:58:01	/CCC.html	http://www.○○○.co.jp/AAA.html	0
View5	Visit1	Visitor1	2002/2/14	18:59:10	/EEE.html	http://www.○○○.co.jp/CCC.html	0
View6	Visit1	Visitor1	2002/2/14	19:00:02	/search.cgi?q=banana	http://www.○○○.co.jp/EEE.html	1
View8	Visit1	Visitor1	2002/2/14	19:01:23	/FFF.html	http://www.○○○.co.jp/search.cgi?q=banana	0
View13	Visit1	Visitor1	2002/2/14	19:02:25	/GGG.html	http://www.○○○.co.jp/FFF.html	0
View2	Visit2	Visitor2	2002/2/14	18:57:30	/BBB.cgi	-	0
View4	Visit2	Visitor2	2002/2/14	18:58:40	/search.cgi?q=food	http://www.○○○.co.jp/BBB.html	1
...

【0537】

始端フラグは「1」がサブアクセスシーケンスの始端を「0」がそれ以外を示す。この際、顧客サイト内検索システム特定情報94と、顧客サイト外検索システム特定情報95を用いる。顧客サイト内検索システム特定情報94の例として上記表5を、顧客サイト外検索システム特定情報95の例として表29を示す。

【0538】

【表29】

表29 顧客サイト外検索システム特定情報

検索システムのURL	クエリーパターン
http://www.□□□.ne.jp/search.cgi?kw=vegetable	kw
...	...

【0539】

例えば上記表28においてView1は、レファラに表29で示した「検索システムのURL」「クエリーパターン」に合致する「http://www.□□□.ne.jp/search.cgi?kw=vegetable」を持っている。このため、「顧客サイト61外の検索システムが出力した検索結果ページをレファラ情報として持つ顧客サイト61内のWebページへのアクセス」すなわちサブアクセスシーケンスの始端となり、表28の始端フラグに「1」が記録される。また、表28においてView6は、リクエストに上記表5で示した「検索システムの場所」「クエリーパターン」に合致する「/search.cgi?q=banana」を持っている。このため、「顧客サイト61内の検索システムによって出力された検索結果ページへのアクセス」すなわちサブアクセスシーケンスの始端となり、表28の始端フラグに「1」が記録される。

【0540】

次に、上記サブアクセスシーケンス抽出規則をもとに、サブアクセスシーケンス終端判定機能93は、表28から終端となるアクセスを判定する。終端とされたものには終端フラグをつける。表30に処理の結果生成された一時的な第3の表を示す。

【0541】

【表30】

表30 一時的な第3の表

VisitID	VisitorID	VisitorID	日付	時刻	リクエスト	レスポンス	始端・終端フラグ
View1	Visit1	Visitor1	2002/2/14	18:56:14	/AAA.html	http://www.□□□.ne.jp/search.cgi?kw=vegitable	1
View3	Visit1	Visitor1	2002/2/14	18:58:01	/CCC.html	http://www.○○○.co.jp/AAA.html	0
View5	Visit1	Visitor1	2002/2/14	18:59:10	/EEE.html	http://www.○○○.co.jp/CCC.html	2
View6	Visit1	Visitor1	2002/2/14	19:00:02	/search.cgi?q=banana	http://www.○○○.co.jp/EEE.html	1
View8	Visit1	Visitor1	2002/2/14	19:01:23	/FFF.html	http://www.○○○.co.jp/search.cgi?q=banana	0
View13	Visit1	Visitor1	2002/2/14	19:02:25	/GGG.html	http://www.○○○.co.jp/FFF.html	2
View2	Visit2	Visitor2	2002/2/14	18:57:30	/BBB.cgi	-	0
View4	Visit2	Visitor2	2002/2/14	18:58:40	/search.cgi?q=food	http://www.○○○.co.jp/BBB.html	1
...

【0542】

始端・終端フラグは「2」がサブアクセスシーケンスの終端を、「1」がサブアクセスシーケンスの始端を「0」がそれ以外を示す。なお、サブアクセスシーケンスの始端と終端が重なっている場合には「3」とする。例えば上記表28において、View5は「サブアクセスシーケンスの始端からみて次のサブアクセスシーケンスの始端の一つ前の顧客サイト内のアクセス」である。このため、サブアクセスシーケンスの終端となり、表30の始端・終端フラグに「2」が記録される。また表28において、View13は「サブアクセスシーケンスの始端からみて次のサブアクセスシーケンスの始端が存在せず、かつ顧客サイト61内のアクセスシーケンスの終端のアクセス」である。このため、サブアクセスシーケンスの終端となり、表30の始端・終端フラグに「2」が記録される。

【0543】

そして、この始端と終端からサブアクセスシーケンスを抽出する。始端・終端フラグが「1」からはじまり「2」で終わる区間、もしくは「3」のものに、重複のないアクセス識別子（SubVisitID）（以下、サブアクセスシーケンス識別子という）を付与する。ただし、同一の区間内には同一のアクセス識別子を付与する。また、「2」からはじまり「1」で終わる区間、すなわち「SubVisitID」に値が存在しない場合は「-」で表される。そして、ViewIDで再び並び替える。

【0544】

表31に、抽出されたサブアクセスシーケンスデータの例を示す。

【0545】

【表 3 1】

表3 1 サブアクセスシーケンスデータ

ViewID	VisitID	VisitorID	SubVisitID	日付	時刻	リクエスト	レフアラ	サイト外検索システムフラグ
View1	Visit1	Visitor1	SubVisit1	2002/2/14	18:56:14	/AAA.html	http://www.□□□.ne.jp/search.cgi?kw=vegetable	1
View2	Visit2	Visitor2	...	2002/2/14	18:57:30	/BBB.cgi	-	-
View3	Visit1	Visitor1	SubVisit1	2002/2/14	18:58:01	/CCC.html	http://www.○○○.co.jp/AAA.html	-
View4	Visit2	Visitor2	SubVisit3	2002/2/14	18:58:40	/search.cgi?q=food	http://www.○○○.co.jp/BBB.html	-
View5	Visit1	Visitor1	SubVisit1	2002/2/14	18:59:10	/EEE.html	http://www.○○○.co.jp/CCC.html	-
View6	Visit1	Visitor1	SubVisit2	2002/2/14	19:00:02	/search.cgi?q=banana	http://www.○○○.co.jp/EEE.html	0
...
View8	Visit1	Visitor1	SubVisit2	2002/2/14	19:01:23	/FFF.html	http://www.○○○.co.jp/search.cgi?fn=banana	-
...
View13	Visit1	Visitor1	SubVisit1	2002/2/14	19:02:25	/GGG.html	http://www.○○○.co.jp/FFF.html	-
...

【0546】

ここで改めて、上記図43を見てみる。図43におけるサブアクセスシーケンス108が上記表31のSubVisit1に、サブアクセスシーケンス109が上記表31のSubVisit2に対応している。表31において、SubVisit1は「AAA.html」→「CCC.html」→「EEE.html」、SubVisit2は「search.cgi?q=food」→「FFF.html」→「GGG.html」となっており、図43のサブアクセスシーケンス108、サブアクセスシーケンス109の画面遷移と合致している。これより、サブアクセスシーケンス判定機能91が、サブアクセスシーケンス抽出規則どおりに機能していることがわかる。

【0547】

上記表31に示すように、サブアクセスシーケンスデータには、サイト外検索システムフラグが設定される。

【0548】

サイト外検索システムフラグは、顧客サイト61外の検索システムによる検索結果ページに基づいて定まるサブアクセスシーケンスであれば「1」、顧客サイト61内の検索システムによる検索結果ページに基づいて定まるサブアクセスシーケンスであれば「0」、それ以外のものに関しては「-」となる。このサイト外検索システムフラグは、上記図38及び図39の説明で述べたように、検索システムが顧客サイト61の内外のいずれにあるかで検索後閲覧時間の算出方法を切り替えるために用いられる。

【0549】

キーワード抽出機能87は、エンコードされた検索キーワードについて判読可能な文字列へのデコードを行い、キーワードデータを生成する。

【0550】

本実施の形態に係るキーワード抽出機能87は、以下の第1及び第2の特徴を持つが、他の点については上記図1で説明したキーワード抽出機能14と同様である。

【0551】

第1に、キーワード抽出機能87は、顧客サイト61外の検索システムについても検索キーワードを抽出する。顧客サイト61外の検索システムに入力された

検索キーワードの抽出は、上記第5の実施の形態で説明したキーワード抽出機能と同様の手法で行うことができる。

【0552】

第2に、キーワード抽出機能87は、顧客サイト61内の検索システムが出力した検索結果ページ2へのアクセスが、顧客サイト61外の検索システムが出力した検索結果ページ2をレファラ情報を持つWebページへのアクセスにも該当する場合に、顧客サイト61内の検索システムが出力した検索結果ページ2へのアクセスに対する検索キーワードを抽出する。

【0553】

キーワード抽出機能87によるキーワードデータの出力形式は、上記表7と同様となる。

【0554】

満足度算出機能89は、サブアクセスシーケンス抽出機能86によって生成されたサブアクセスシーケンスデータに基づいて、検索後閲覧時間を計算し、閲覧者の満足度を算出する。

【0555】

満足度算出機能89は、以下の第1及び第2の特徴を持つが、他の点については上記図1で説明した満足度算出機能15と同様である。

【0556】

第1に、満足度算出機能89は、質的評価規則における検索後閲覧時間を、アクセスシーケンスごとではなく、サブアクセスシーケンスごとに求め、満足度を算出する。

【0557】

第2に、満足度算出機能89は、上記図38及び図39の説明で述べたように、検索システムが顧客サイト61の内外のいずれにあるかで検索後閲覧時間の算出方法を切り替え、検索後閲覧時間を求める。サブアクセスシーケンスが、顧客サイト61内外のいずれの検索システムによるかの判断は、上記表27のサイト外検索システムフラグに基づいて判断する。

【0558】

表32に、抽出された満足度算出データの例を示す。

【0559】

【表32】

表32 満足度算出データ

ViewID	SubVisitID	満足度
View1	SubVisit1	40
View2	-	-
View3	SubVisit1	40
View4	SubVisit3	10
View5	SubVisit1	40
...

【0560】

表32には、各アクセス識別子と、そのアクセスの属するサブアクセスシーケンスを示すサブアクセスシーケンス識別子と、そのサブアクセスシーケンス識別子に対応する満足度とを関係付けたレコードが含まれている。項目「SubVisitID」においてサブアクセスシーケンス識別子が存在しない部分には、満足度も存在しないため「-」が表記される。

【0561】

満足度・キーワード関連付け機能16は、キーワード抽出機能87によって生成されたキーワードデータと、表28に示す満足度算出データとに基づいて、検索キーワードとその検索キーワードに対応するサブアクセスシーケンスの満足度との関係付けを行い、上記表11のような形式の関係データを算出する。

【0562】

図44は、本実施の形態に係る検索キーワード分析システム83の動作の一例を示すフローチャートである。

【0563】

ステップZ1において、ログ前処理機能12は、Webアクセスログ7に含まれる情報から必要な情報を入力し、前処理データを作成する。

【0564】

ステップZ2において、サブアクセスシーケンス抽出機能86は、前処理データとアクセスシーケンス抽出規則とに基づいて、アクセスシーケンスを抽出し、

アクセスシーケンスデータを作成する。

【0565】

ステップZ3において、サブアクセスシーケンス抽出機能86は、アクセスシーケンスデータとサブアクセスシーケンス抽出規則とに基づいて、サブアクセスシーケンスを抽出し、サブアクセスシーケンスデータを作成する。

【0566】

ステップZ4において、満足度算出機能15は、サブアクセスシーケンスデータと満足度算出規則とに基づいて、サブアクセスシーケンス毎の満足度を算出し、満足度算出データを作成する。

【0567】

ステップZ5において、キーワード抽出機能14は、前処理データに基づいて、顧客サイト6内の検索システム及び顧客サイト6外の検索システムに入力された検索キーワードを抽出し、キーワードデータを作成する。

【0568】

ステップZ6において、満足度・キーワード関連付け機能16は、満足度算出データと、キーワードデータに基づいて、検索キーワードごとの満足度を算出し、関係データを作成する。

【0569】

なお、ステップZ5は、ステップZ2～Z4の前に実行されてもよく、後に実行されてもよく、並列に実行されてもよい。

【0570】

以上説明した本実施の形態においては、一連のアクセスシーケンスにおいて複数の検索キーワードが入力された場合（複数の検索行為がなされた場合）であっても、複数の検索キーワード毎に質的評価値が求められ、さらに複数の検索キーワード毎に質的評価値と量的評価値を算出できる。

【0571】

一般に、サイト閲覧者による検索システムの利用機会は多く、同じサイト閲覧者による一連のアクセスにおいても複数回検索キーワードが入力されることが多い。本実施の形態では、このような場合であっても、検索キーワード毎に質的評

価値を求めることができ、検索キーワード及びその検索キーワードに基づく検索を実行した検索システムについて十分な精度で分析を行うことができる。

【0572】

なお、本実施の形態においては、質的評価値を示す指標が満足度の場合について説明しているが、質的評価値として成功度などの別の指標を用いてもよい。

【0573】

また、検索後閲覧時間には、本実施の形態で定義した方式以外の方式を適用できる。例えば検索後閲覧時間を、Webページのコンテンツの内容と関連付けて決定するとしてもよい。

【0574】

また、本実施の形態においては、サブアクセスシーケンス抽出機能86において、アクセスシーケンスの抽出後サブアクセスシーケンスを抽出したが、アクセスシーケンスを抽出することなく直接的にサブアクセスシーケンスを抽出してもよい。

【0575】

上記各実施の形態に係る検索キーワード分析システムの各構成要素は、同様の作用を実現可能であれば配置を変更させてもよく、また各構成要素を自由に組み合わせてもよい。

【0576】

また、上記各実施の形態において、検索キーワード分析システムは複数の計算機により構成され、検索キーワード分析プログラムは複数の計算機に分散して配置され、互いに連携を取りつつ処理を実行するとしてもよい。

【0577】

上記各実施の形態に係る検索キーワード分析プログラムは、例えば磁気ディスク（フレキシブルディスク、ハードディスク等）、光ディスク（CD-ROM、DVD等）、半導体メモリなどの記録媒体に書き込んでコンピュータに適用可能である。また検索キーワード分析プログラムは、通信媒体により伝送してコンピュータに適用することも可能である。コンピュータは、記録媒体に記録された検索キーワード分析プログラムを読み込み、検索キーワード分析プログラムによつ

て動作が制御されることにより、上記の機能を実現する。

【0578】

【発明の効果】

以上詳記したように本発明においては、閲覧者の検索結果に対する質的評価値を迅速かつ高精度に求めることができ、この質的評価値に基づいてWebページ又はWebサイトのデザインやコンテンツ、検索システムを改善することができる。

【図面の簡単な説明】

【図1】 本発明の第1の実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図。

【図2】 同実施の形態に係る顧客サイトの構成の一例を示すブロック図。

【図3】 同実施の形態に係る検索キーワード分析システムの動作の一例を示すフローチャート。

【図4】 同実施の形態に係るキーワード抽出機能の動作の一例を示すフローチャート。

【図5】 同実施の形態に係る満足度算出機能の動作の一例を示すフローチャート。

【図6】 本発明の第2の実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図。

【図7】 本発明の第3の実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図。

【図8】 同実施の形態において表示されるグラフの一例を示す図。

【図9】 本発明の第4の実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図。

【図10】 本発明の第5の実施の形態に係る顧客サイトの構成の一例を示す図。

【図11】 サイト外検索システムから顧客サイトへの閲覧遷移の一例を示すブロック図。

【図12】 顧客サイト内の検索キーワード入力画面からサイト外検索シス

テムの検索結果ページへ移り、上記検索結果ページから顧客サイト内のキーワード関連ページに遷移する一例を示すブロック図。

【図13】 本発明の第6の実施の形態に係る顧客サイトの構成の一例を示す図。

【図14】 本発明の第7の実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図。

【図15】 同実施の形態に係る検索キーワード分析システムの動作の一例を示すフローチャート。

【図16】 分析対象のWebサイトをアクセスするために入力された検索キーワードのニーズ値と満足度を計算した結果の一例を示す図。

【図17】 検索キーワード単位のニーズ値と満足度とカテゴリ単位のニーズ値と満足度の関係の一例を示す図。

【図18】 検索キーワードの遷移状況にメッセージを付した状態の一例を示す図。

【図19】 カテゴリの遷移状況にメッセージを付した状態の一例を示す図。

【図20】 カテゴリ形成機能の動作の一例を示すフローチャート。

【図21】 カテゴリ形成機能によるデータの変換結果の一例を示す図

【図22】 同実施の形態に係るトレンド分析機能の動作の一例を示すフローチャート。

【図23】 ニーズ境界値と満足度境界値とによって決定される領域の名称の一例を示す図。

【図24】 同実施の形態に係るトレンド分析機能によって作成されるトレンドデータの一例を示す図

【図25】 同実施の形態において領域が移動した場合に付与されるメッセージの他の例を示す図。

【図26】 同実施の形態に係る出力処理機能によって表示されるグラフの一例を示す図。

【図27】 本発明の第8の実施の形態に係るトレンド分析機能の構成の一

例を示すブロック図。

【図 28】 メッセージと移動方向の関係の具体例を示す図。

【図 29】 同実施の形態に係るトレンド分析機能によって作成されたトレンドデータの一例を示す図。

【図 30】 同実施の形態に係る検索キーワード分析システムの動作の一例を示すフローチャート。

【図 31】 同実施の形態に係る出力処理機能によって出力されたトレンドデータの一例を示す図。

【図 32】 顧客サイトに検索システムが具備される場合のアクセスシーケンスの状態例を示す図。

【図 33】 検索システムが顧客サイト外に存在する場合のアクセスシーケンスの状態例を示す図。

【図 34】 検索システムが顧客サイト外にあり、顧客サイト内のアクセスシーケンスの途中に検索結果ページが存在する場合のアクセスシーケンスの状態例を示す図。

【図 35】 顧客サイト内に検索システムが具備される場合のアクセスシーケンスにおいて複数の検索が行われた状態例を示す図。

【図 36】 検索システムが顧客サイト外に存在する場合のアクセスシーケンスにおいて複数の検索が行われた状態例を示す図。

【図 37】 顧客サイト内の検索システムと顧客サイト外の検索システムの双方を用いてアクセスシーケンスの途中で検索が行われた状態例を示す図。

【図 38】 顧客サイト内の検索システムによる検索結果ページに基づいて抽出されたサブアクセスシーケンスと検索後閲覧時間との関係の一例を示す図。

【図 39】 顧客サイト外の検索システムによる検索結果ページに基づいて抽出されたサブアクセスシーケンスと検索後閲覧時間との関係の一例を示す図。

【図 40】 本発明の第9の実施の形態に係る検索キーワード分析システムの構成の一例を示すブロック図。

【図 41】 同実施の形態に係るサブアクセスシーケンス抽出機能の具体例を示すブロック図。

【図42】 同実施の形態に係るサブアクセスシーケンス判定機能の具体例を示すブロック図。

【図43】 アクセスシーケンスの画面遷移の一例を示す図。

【図44】 同実施の形態に係る検索キーワード分析システムの動作の一例を示すフローチャート。

【図45】 検索システムを用いた代表的なWebページの遷移の例を示す図。

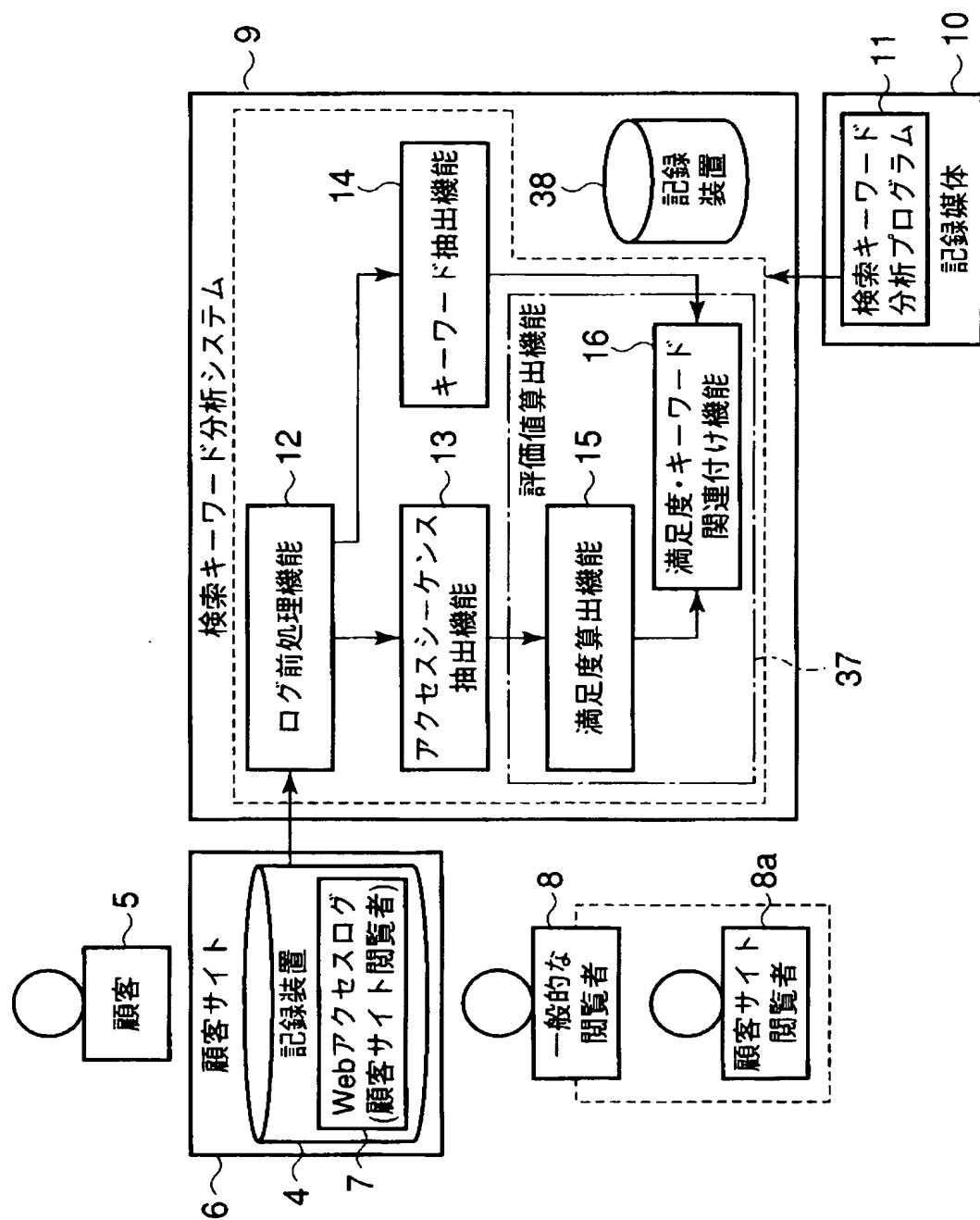
【符号の説明】

4, 38…記録装置、5…顧客、6…顧客サイト、6a…サイト内検索システム、7…Webアクセスログ、8…一般的な閲覧者、8a…顧客サイト閲覧者、9, 18, 24, 29, 42…検索キーワード分析システム、10, 19, 25, 30, 43…記録媒体、11, 20, 26, 31, 44, 85…検索キーワード分析プログラム、12, 12a…ログ前処理機能、13, 13a…アクセスシーケンス抽出機能、14, 14a, 45, 87…キーワード抽出機能、15, 15a, 89…満足度算出機能、16…満足度・キーワード関連付け機能、17, 23, 28, 33…出力データ、21, 21a…ニーズ値算出機能、22, 22a…満足度・ニーズ値関連付け機能、27…グラフ化機能、32…成功度算出機能、341～34n…サイト外検索システム、37, 39, 39a, 40, 46, 88…評価値算出機能、41…成功度・キーワード関連付け機能、52…カテゴリ形成機能、53, 79…トレンド分析機能、54…出力処理機能、86…サブアクセスシーケンス抽出機能

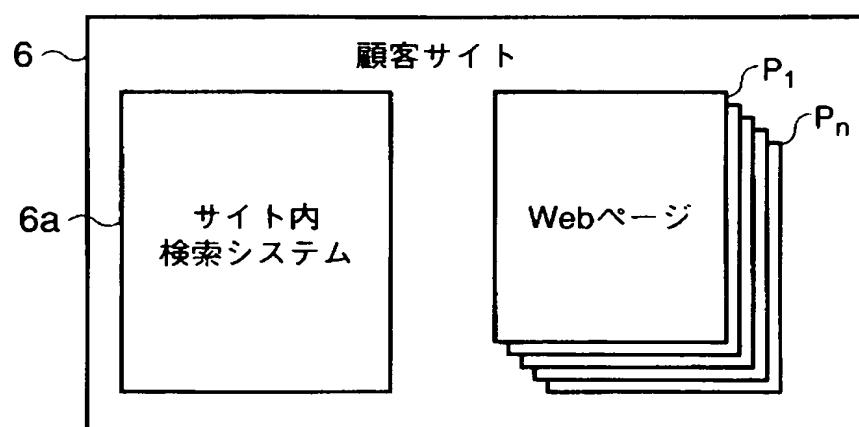
【書類名】

図面

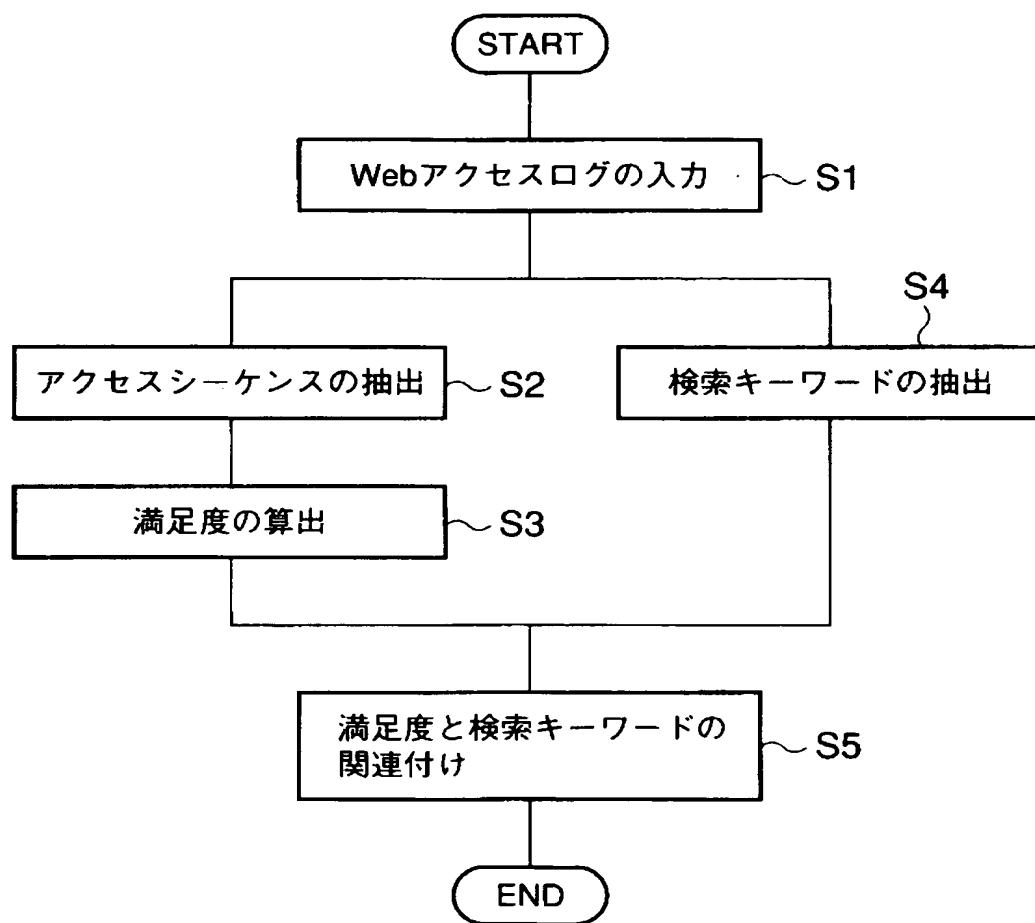
【図 1】



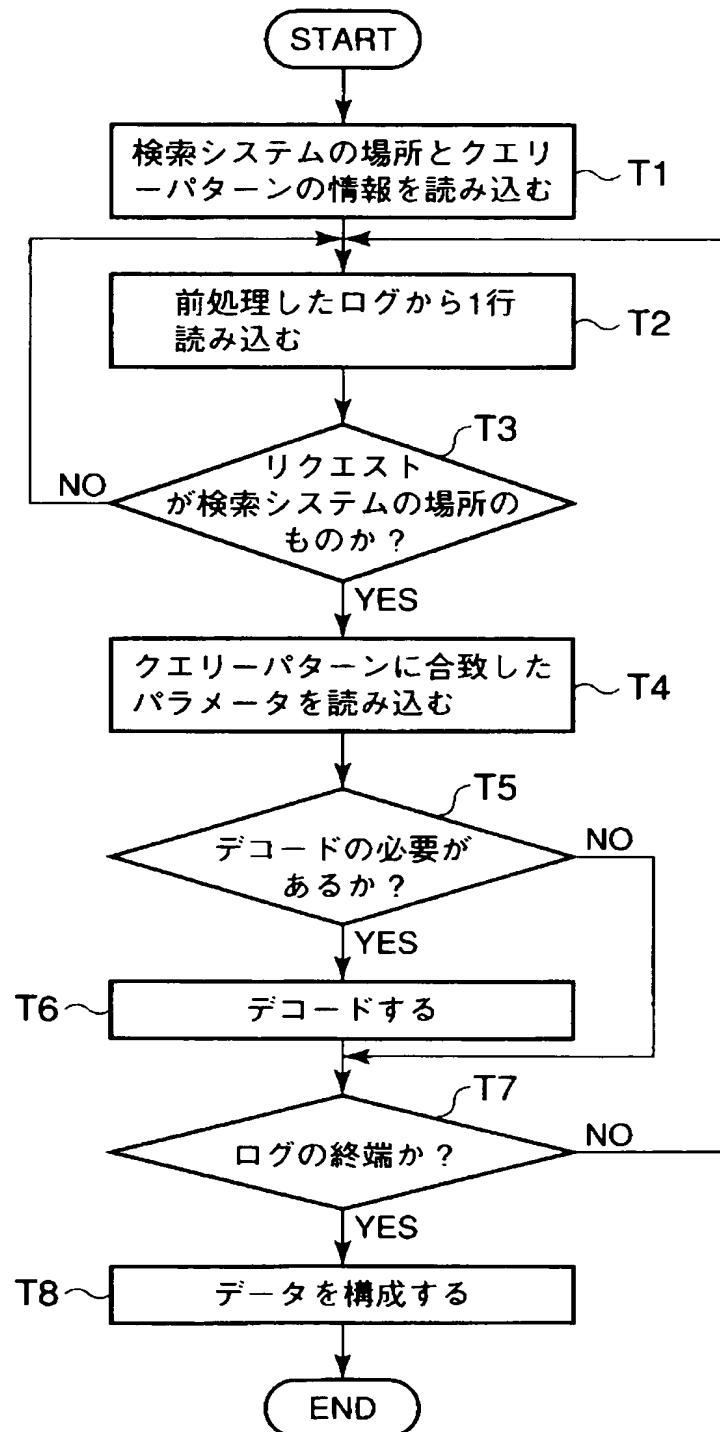
【図2】



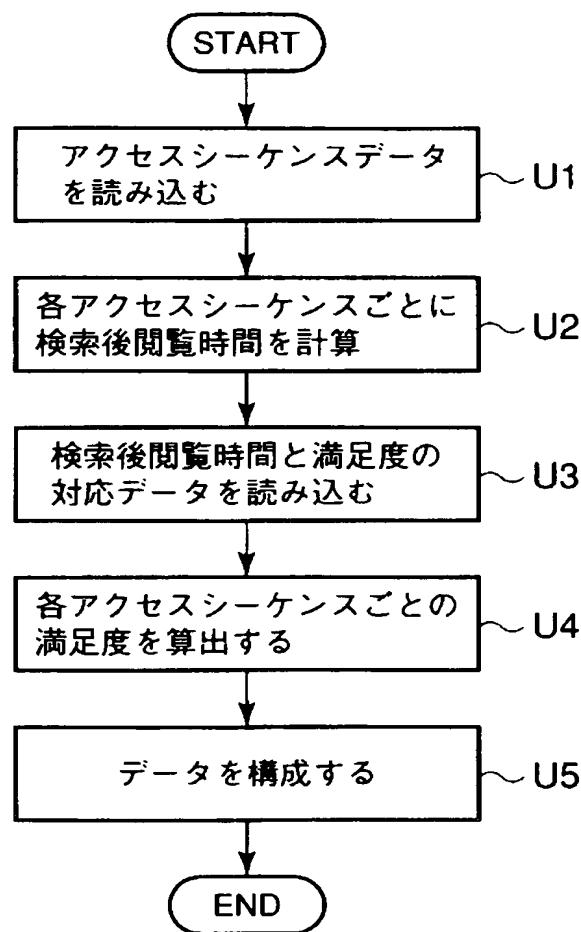
【図3】



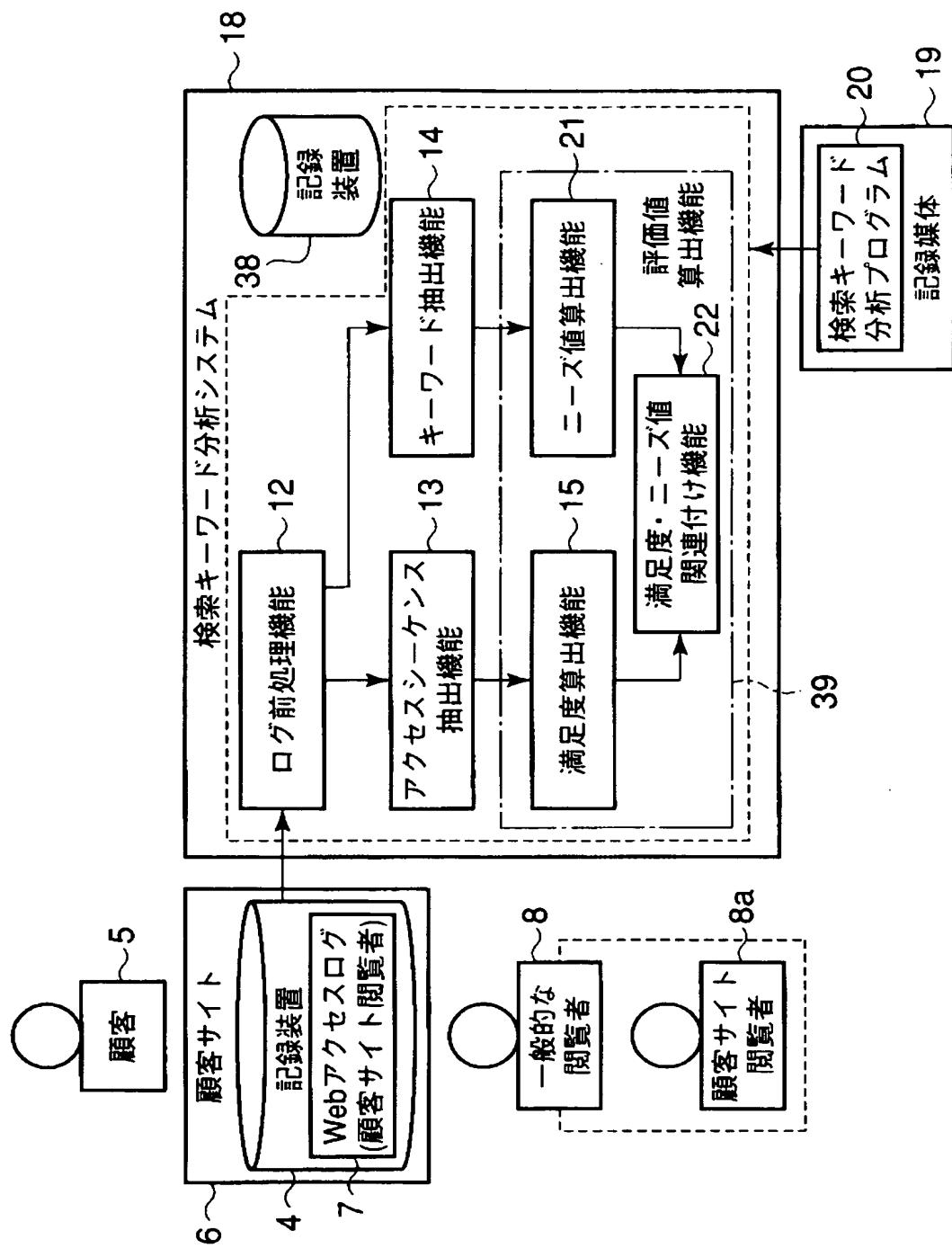
【図4】



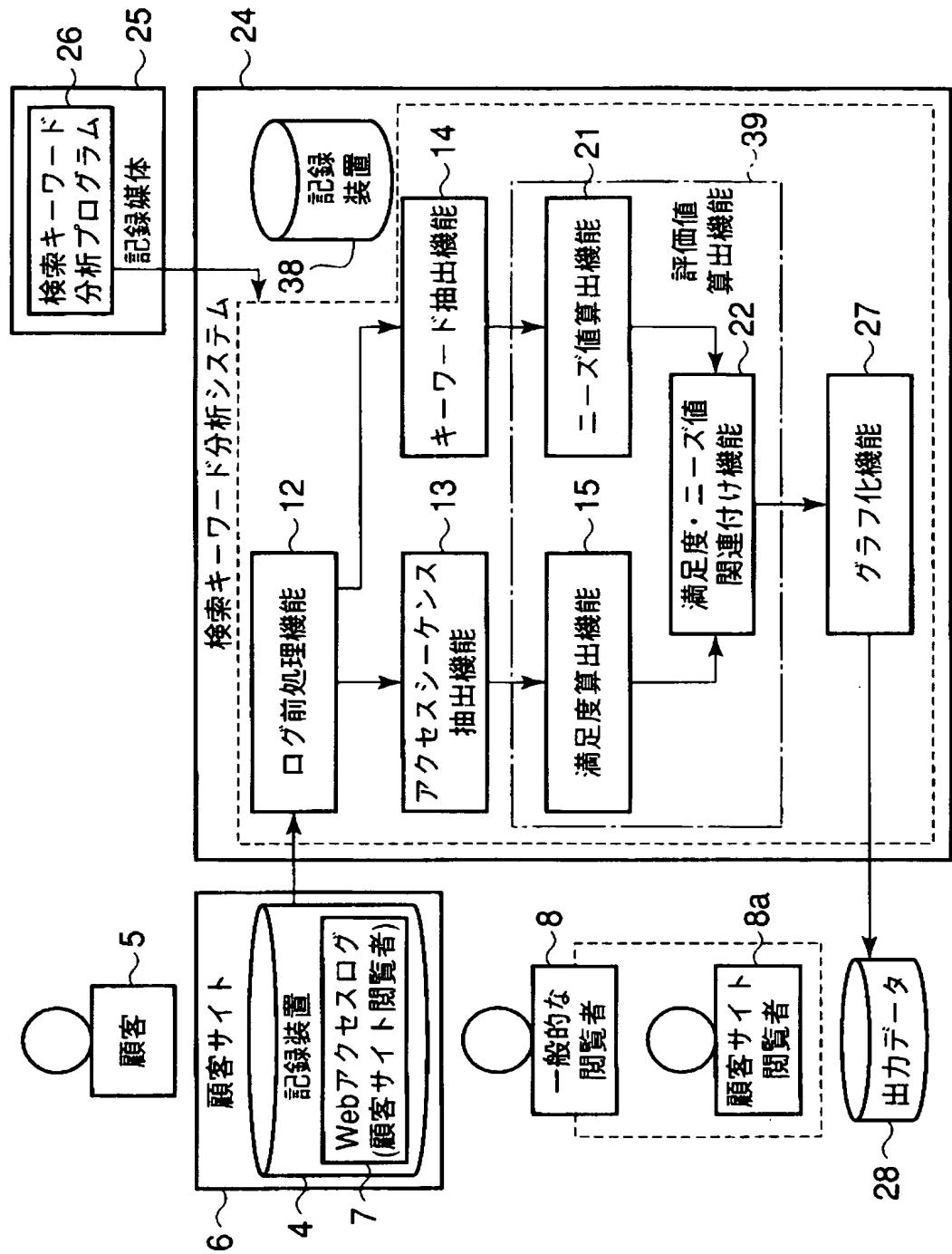
【図5】



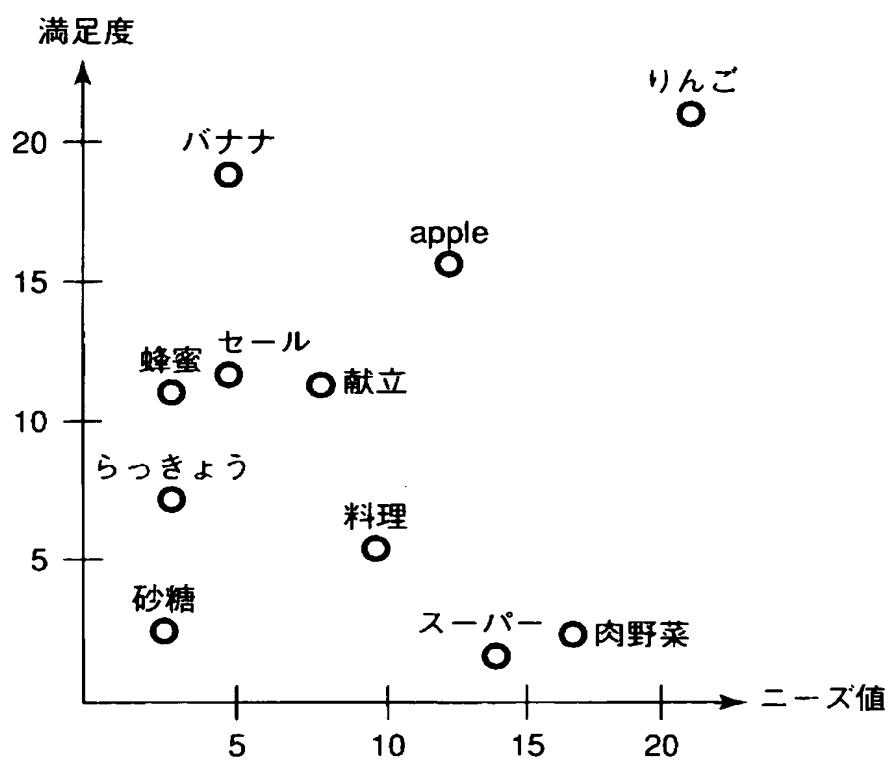
【図 6】



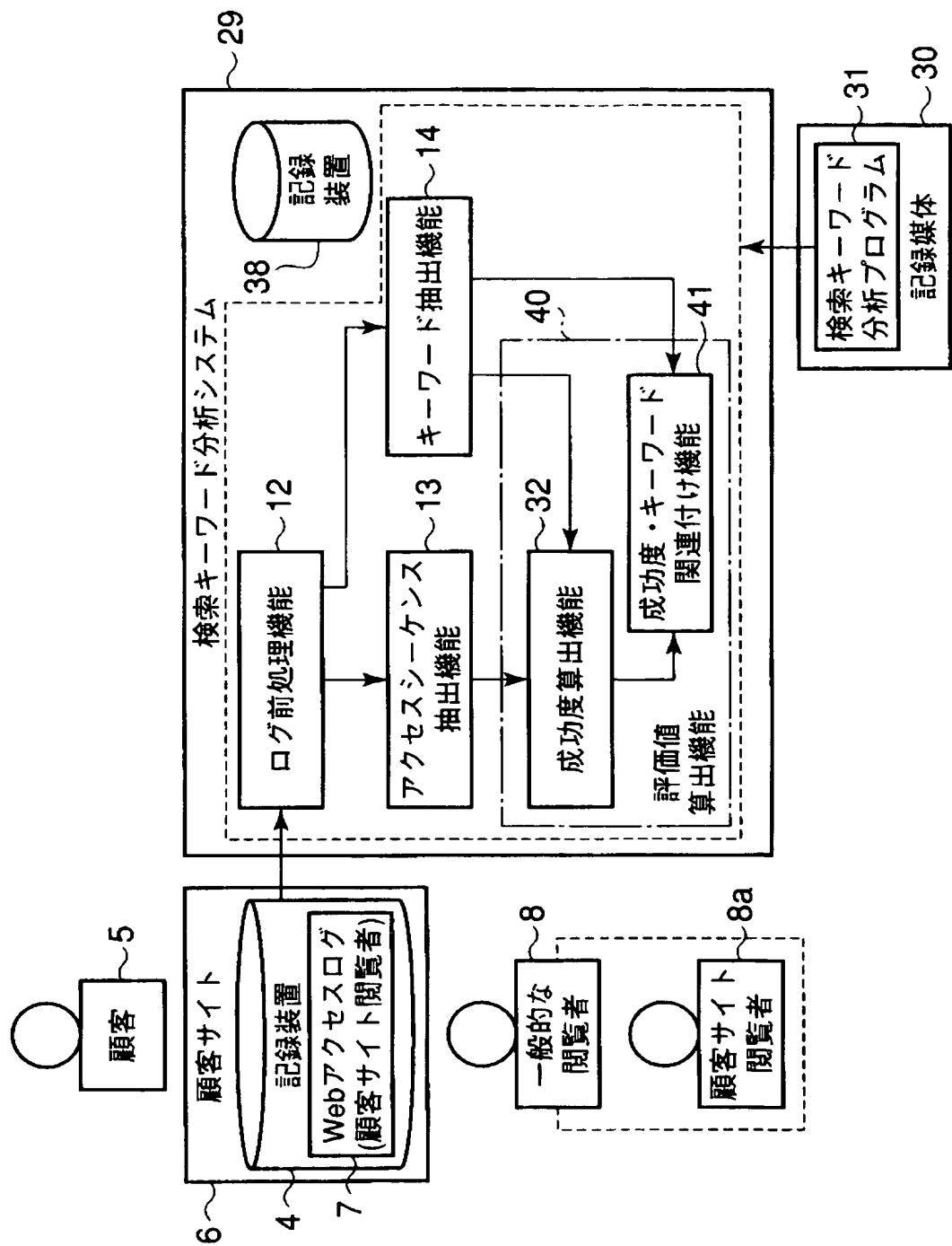
【図 7】



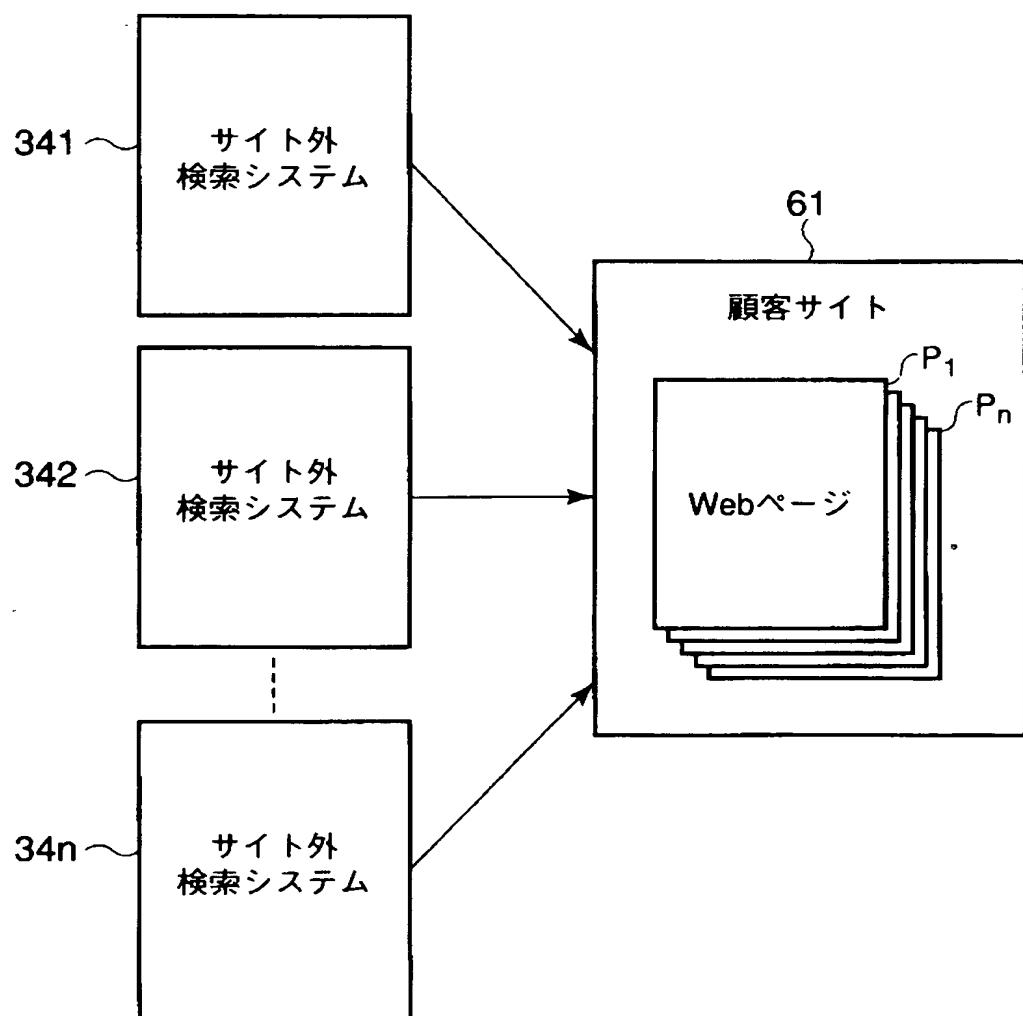
【図 8】



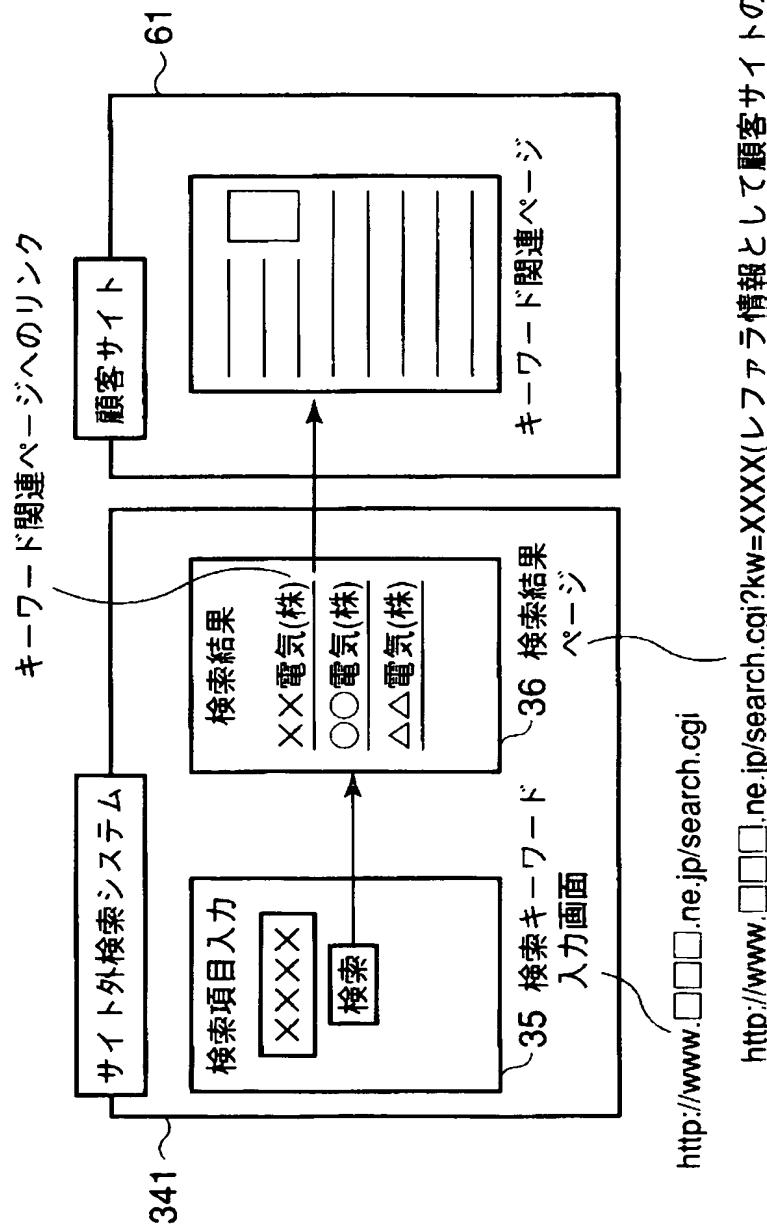
【図9】



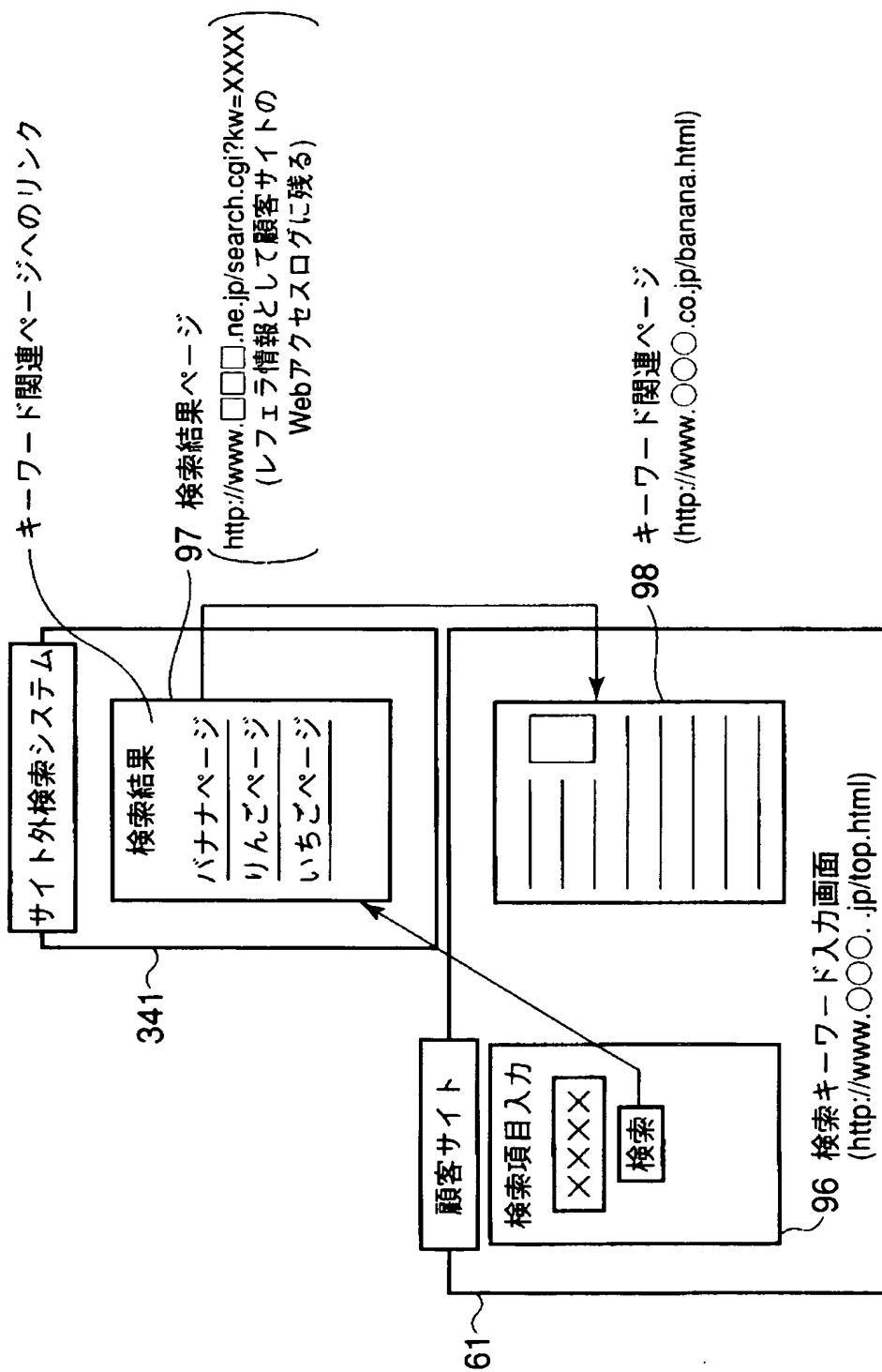
【図10】



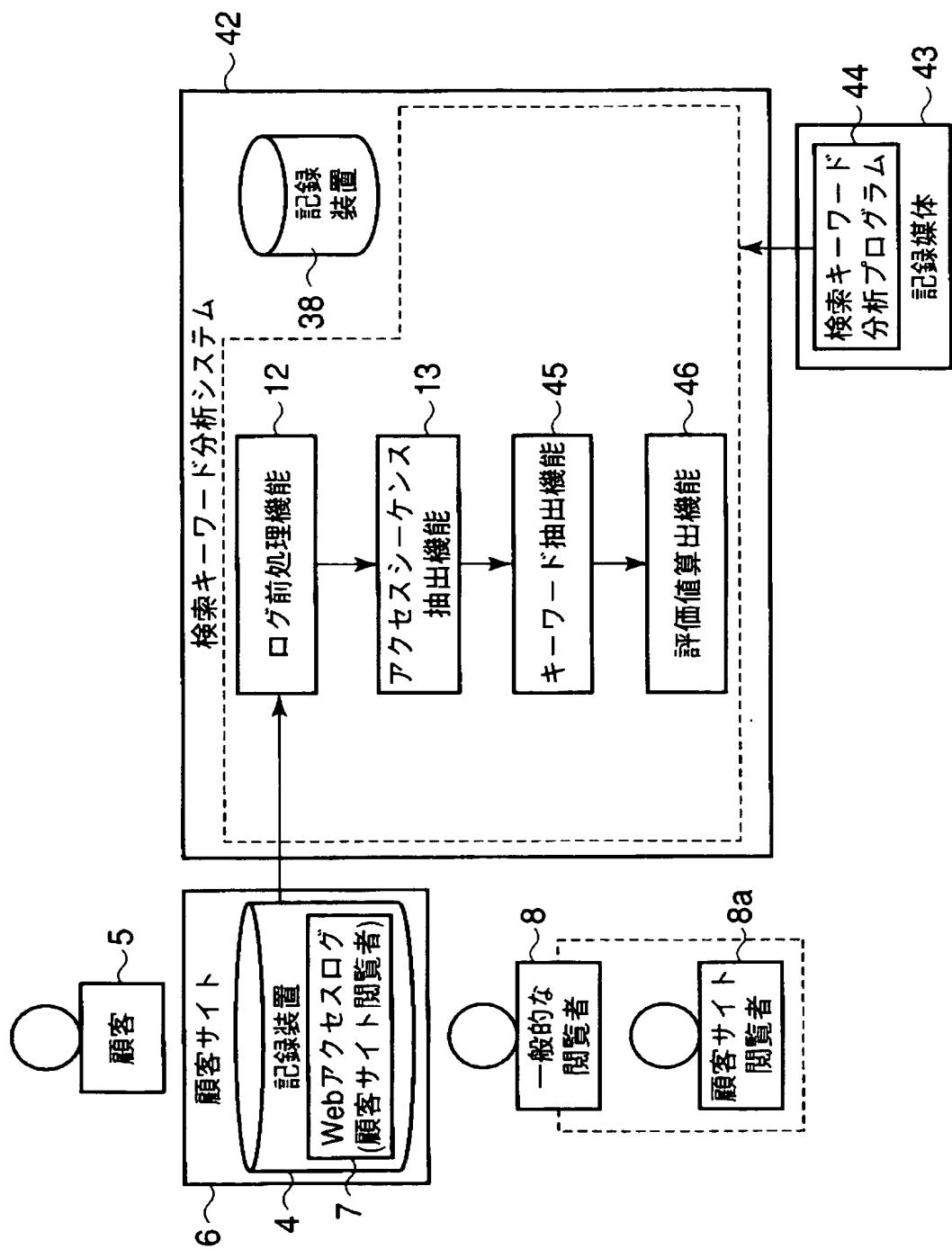
【図11】



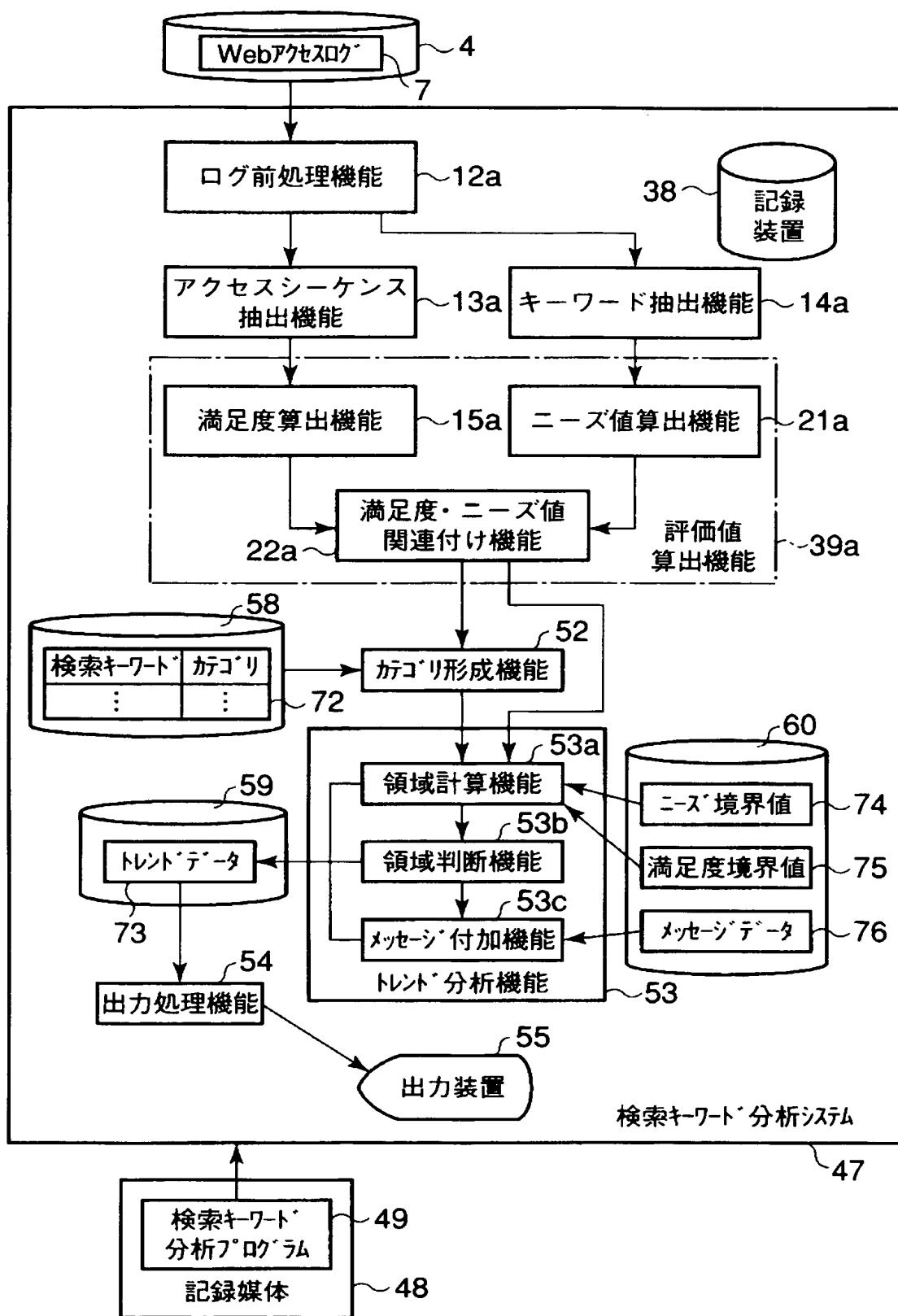
【図12】



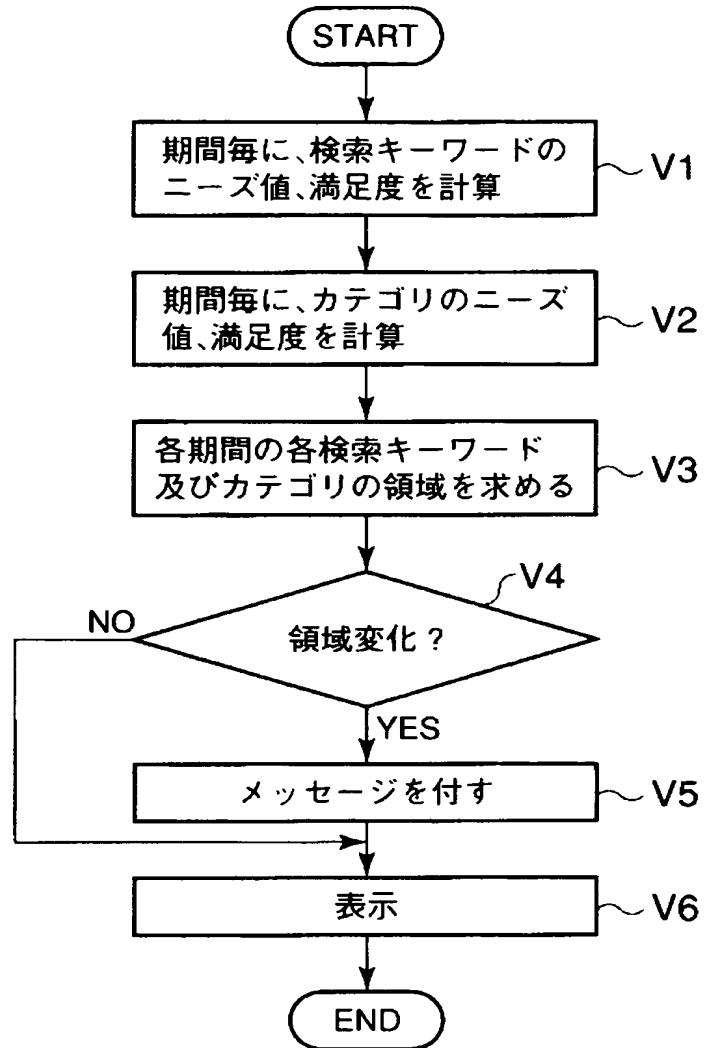
【図13】



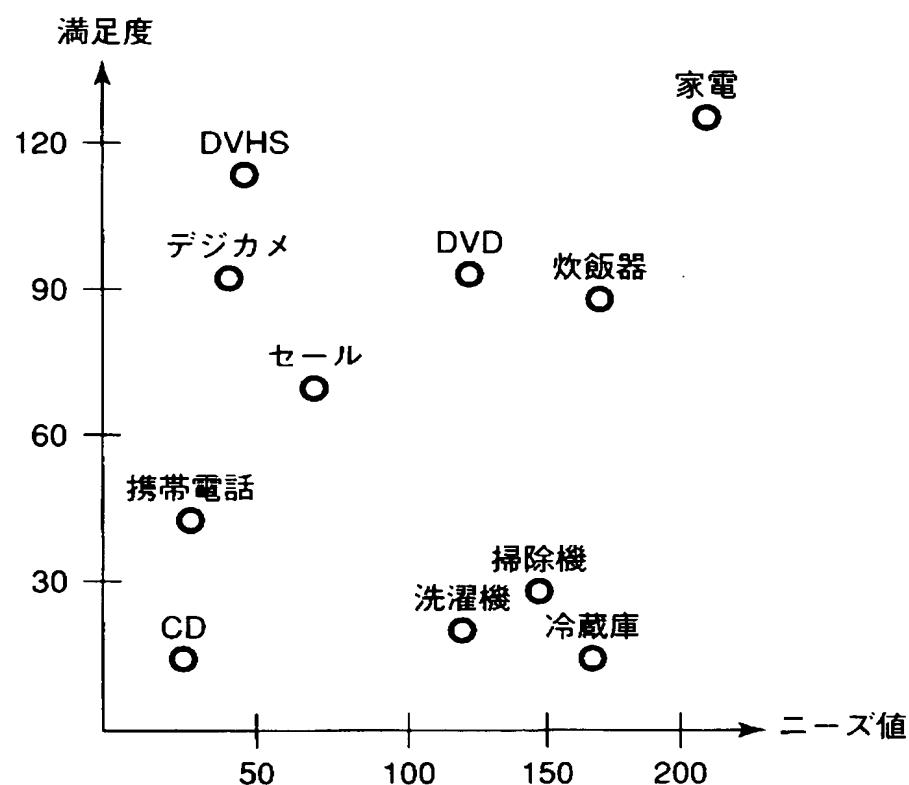
【図14】



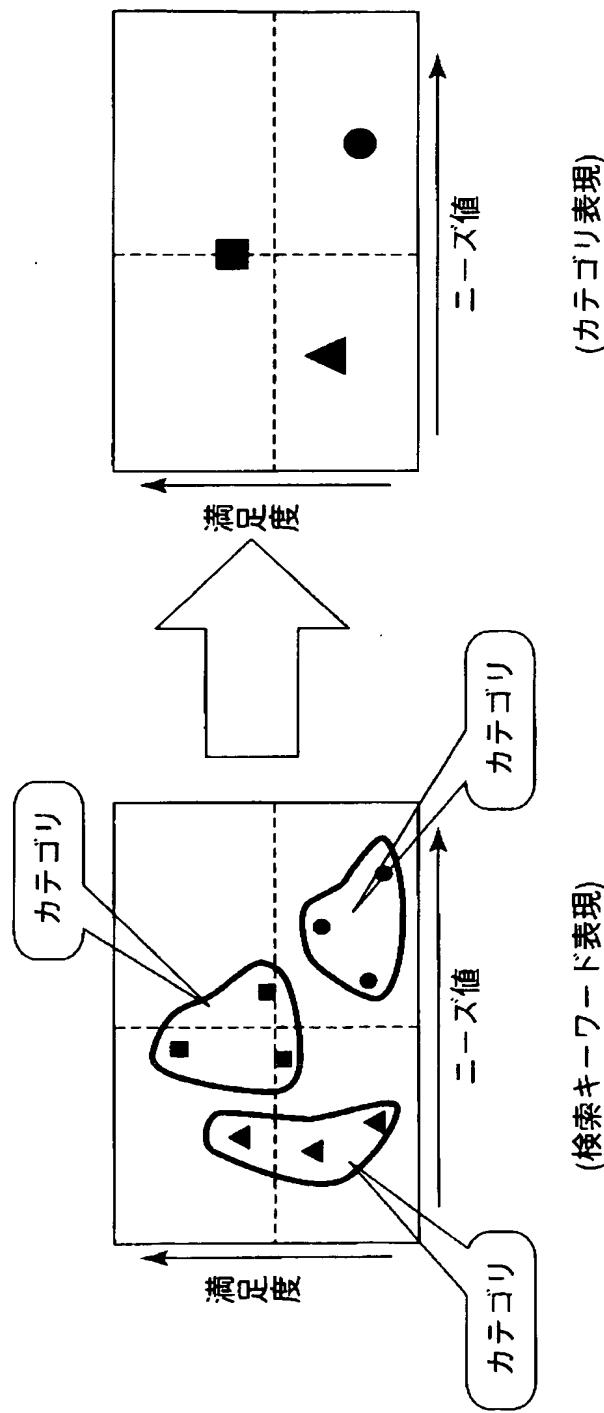
【図15】



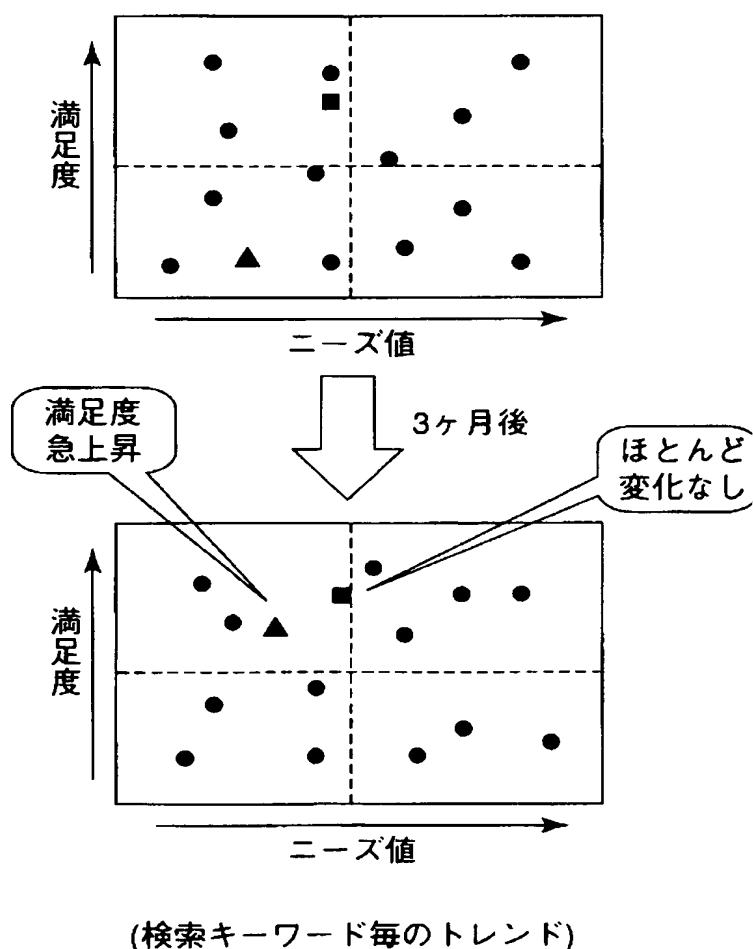
【図16】



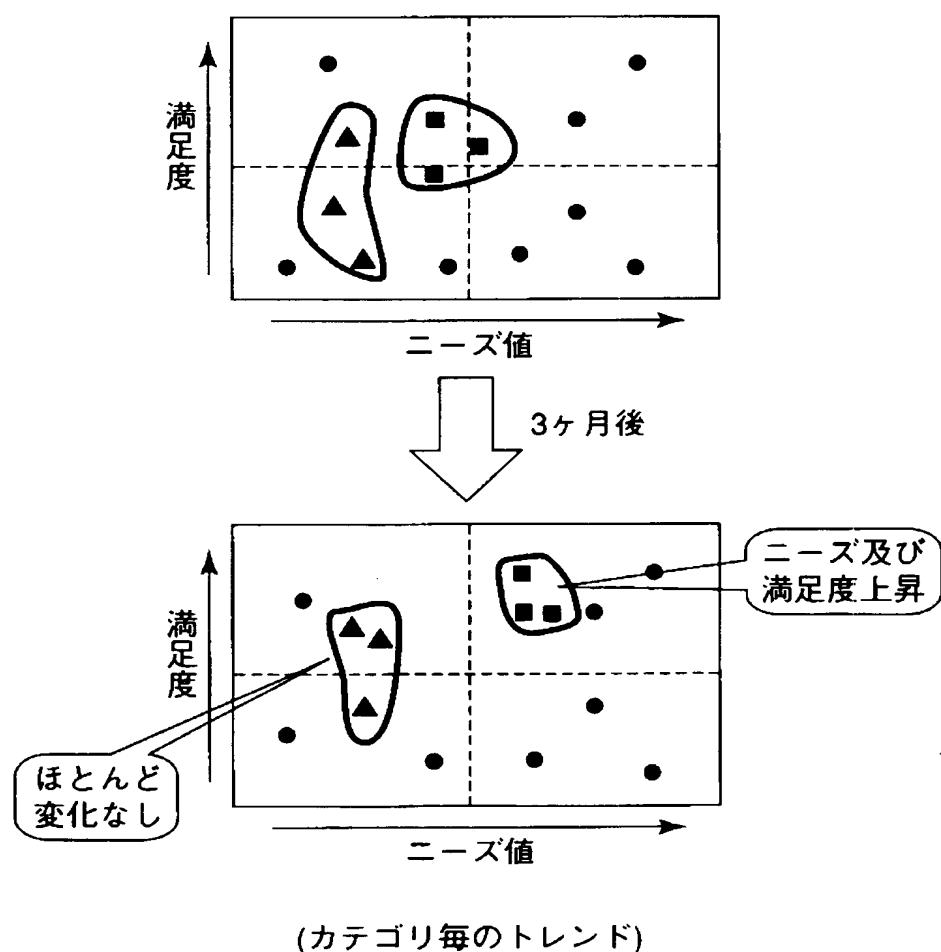
【図17】



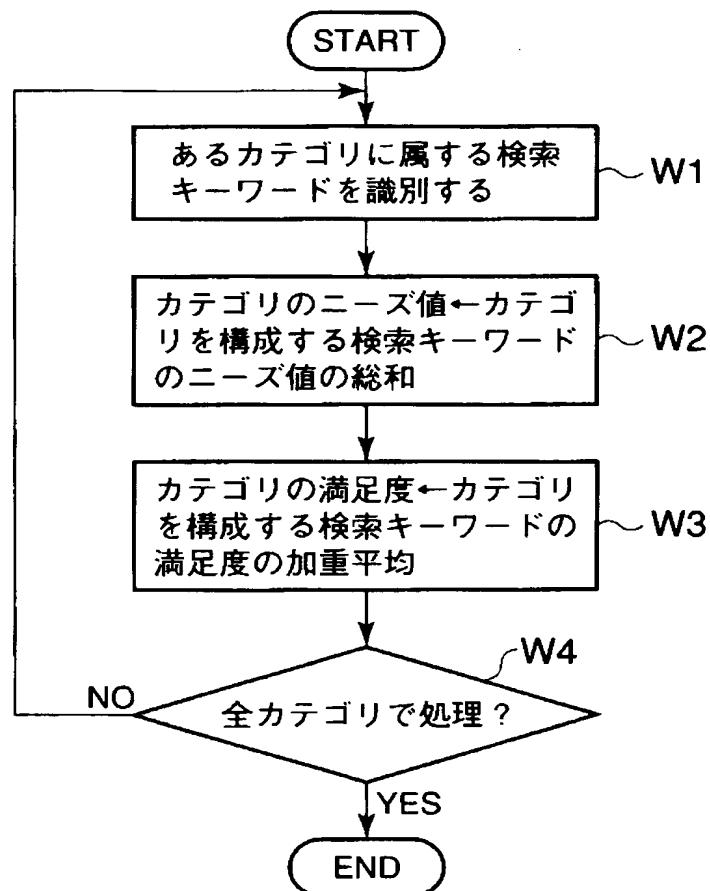
【図18】



【図19】



【図20】



【図21】

72 カテゴリデータ

キーワード	カテゴリ
食べ放題	グルメ
宴会	グルメ
ランチ	グルメ
ぜんそく	健康
ハウスダスト	健康

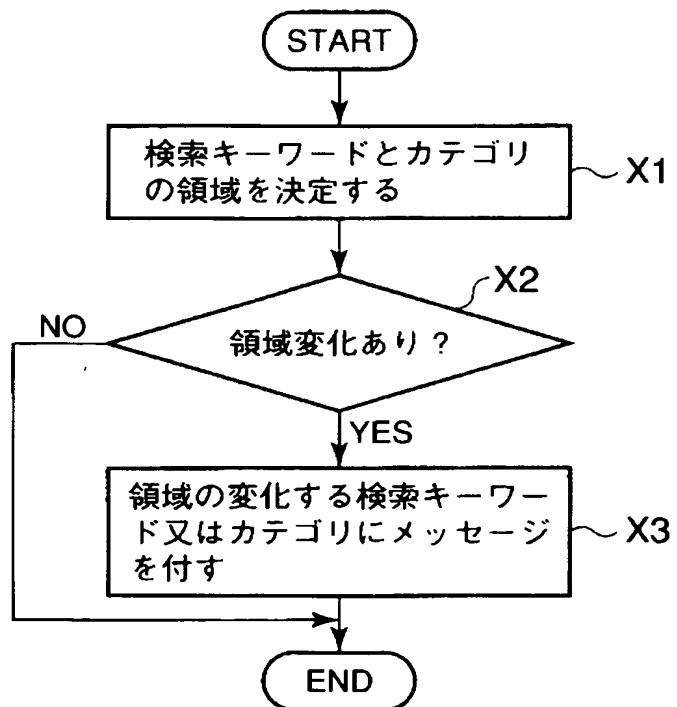
77 関係データ

キーワード	ニーズ	満足度
食べ放題	155	169
宴会	93	638
ランチ	124	221
ぜんそく	77	226
ハウスダスト	177	106

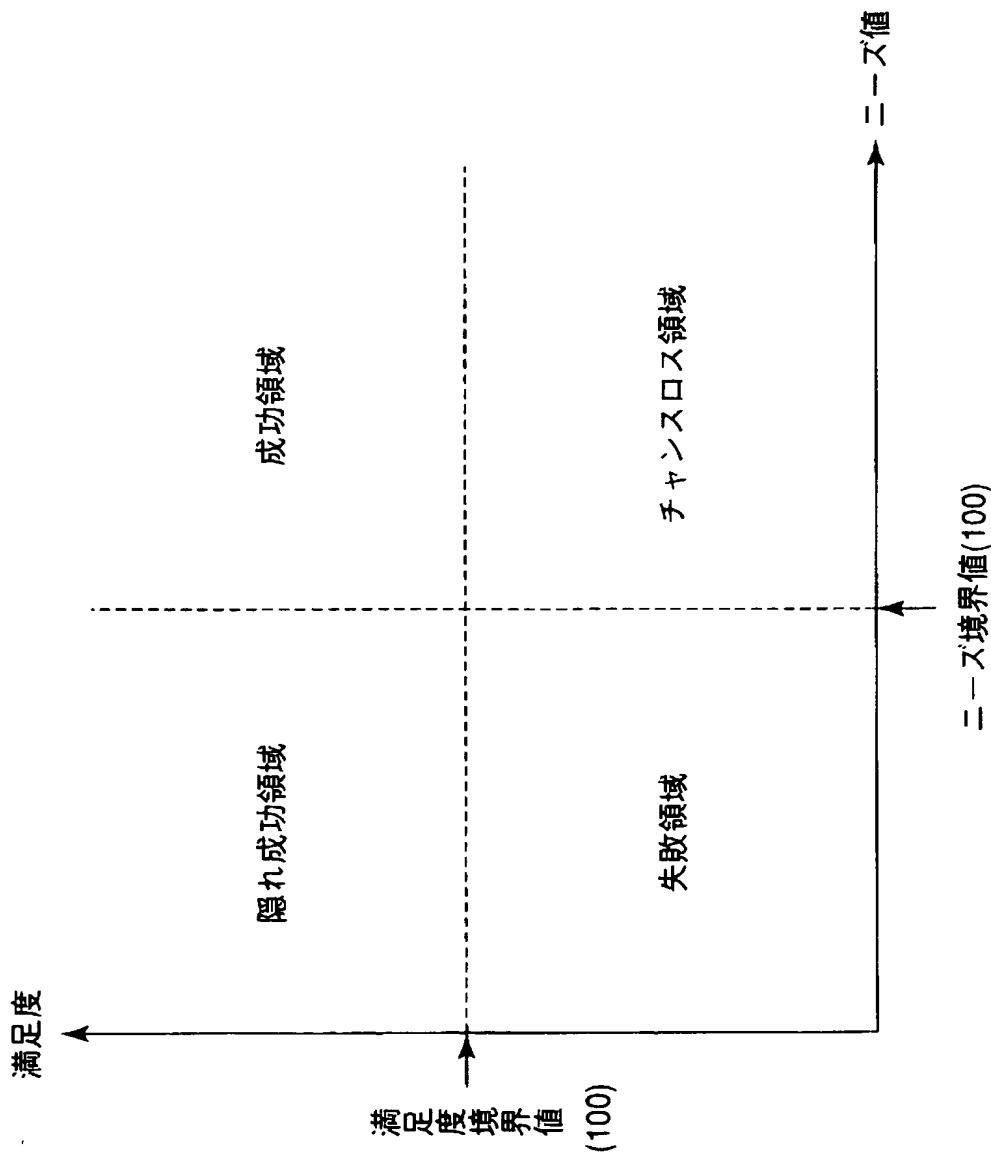
78 カテゴライズデータ

カテゴリ	ニーズ	満足度
グルメ	372	304
健康	254	142

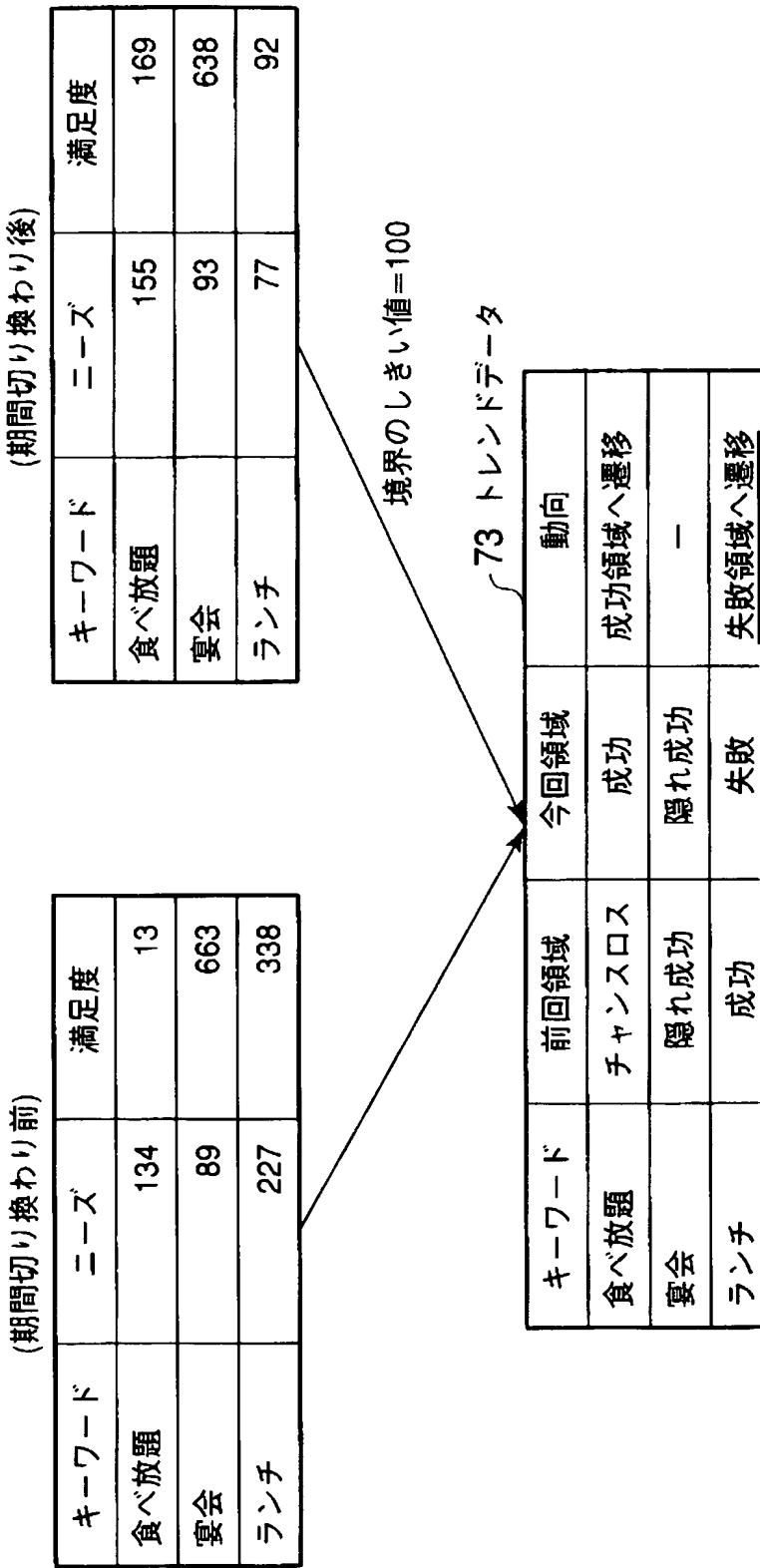
【図22】



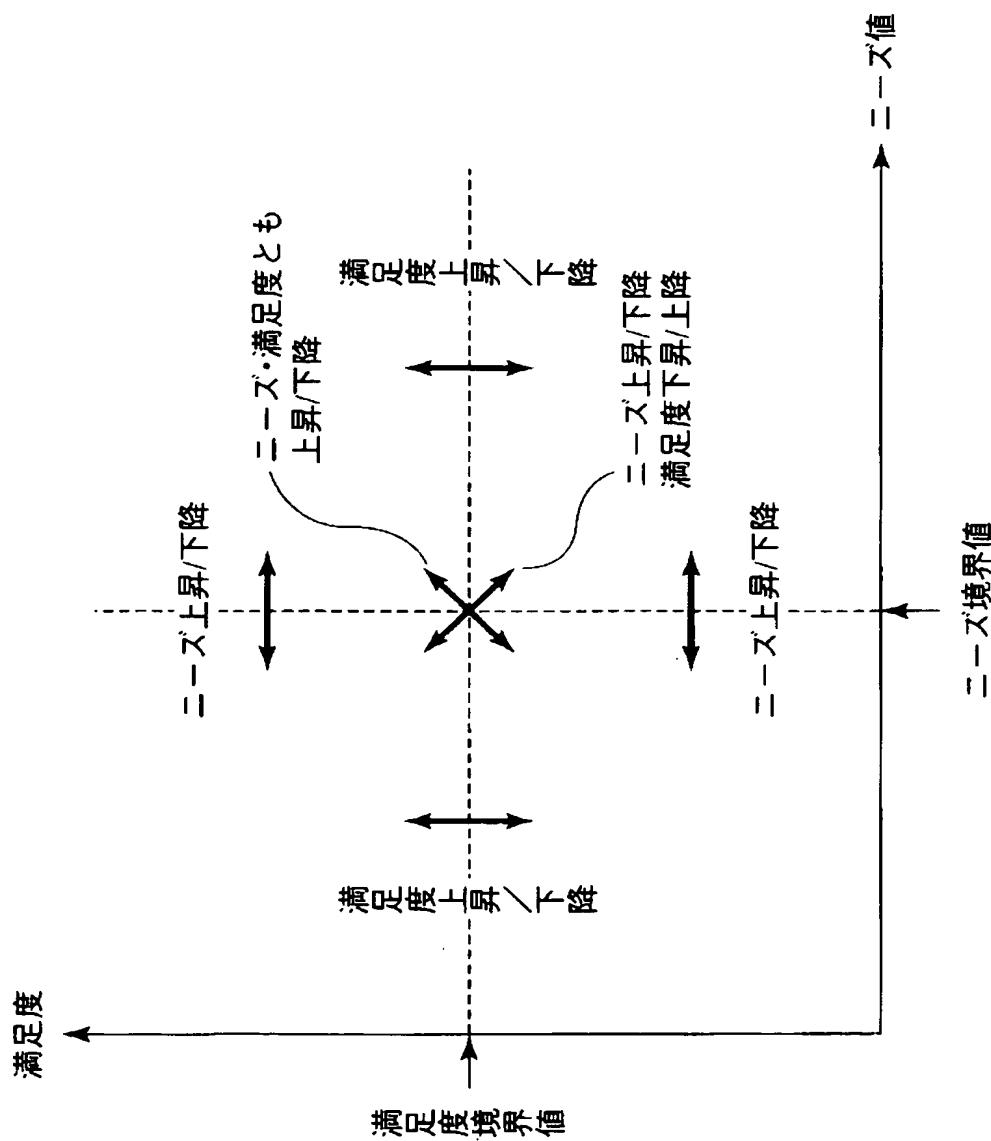
【図23】



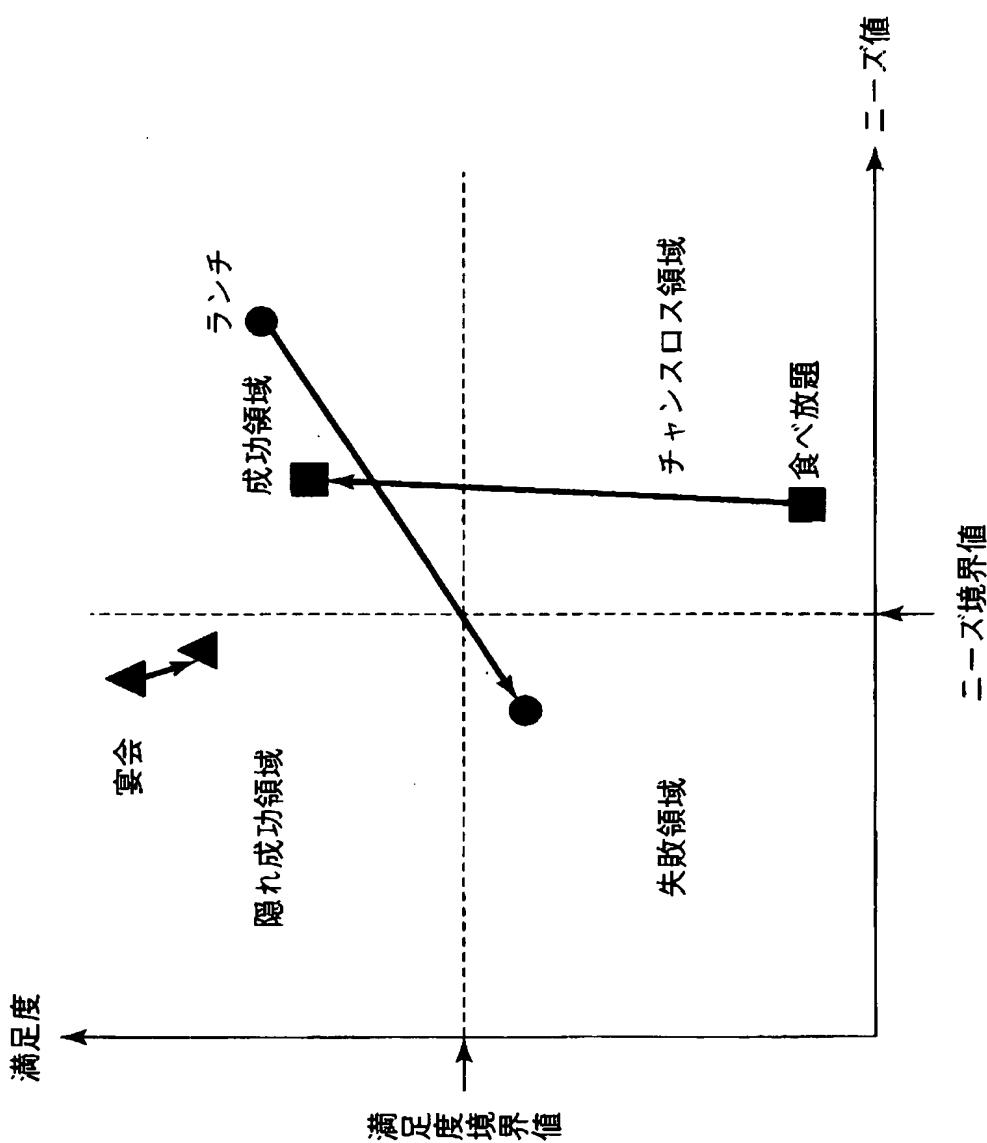
【図24】



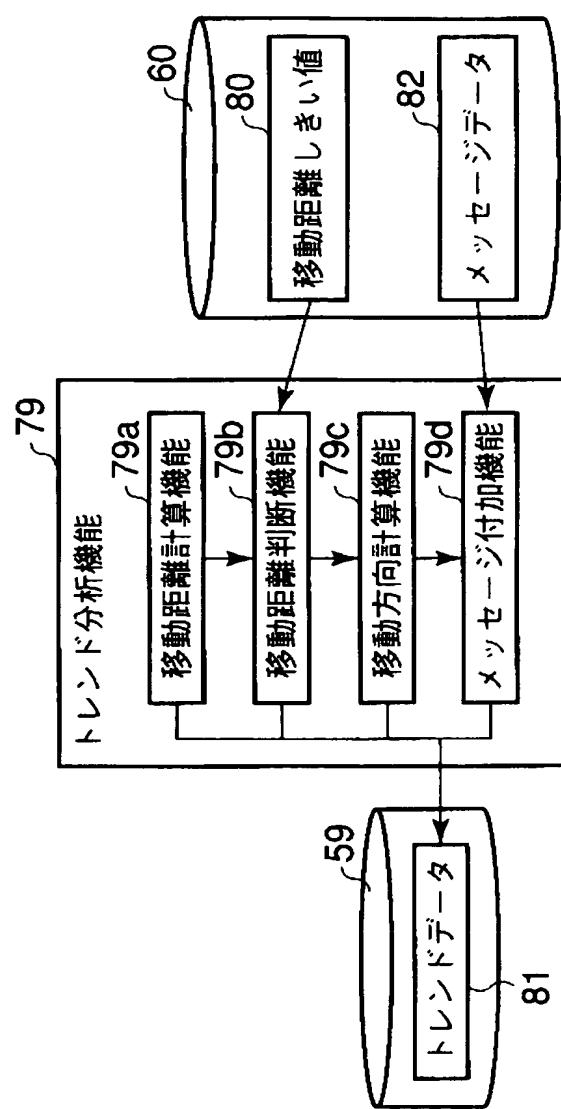
【図25】



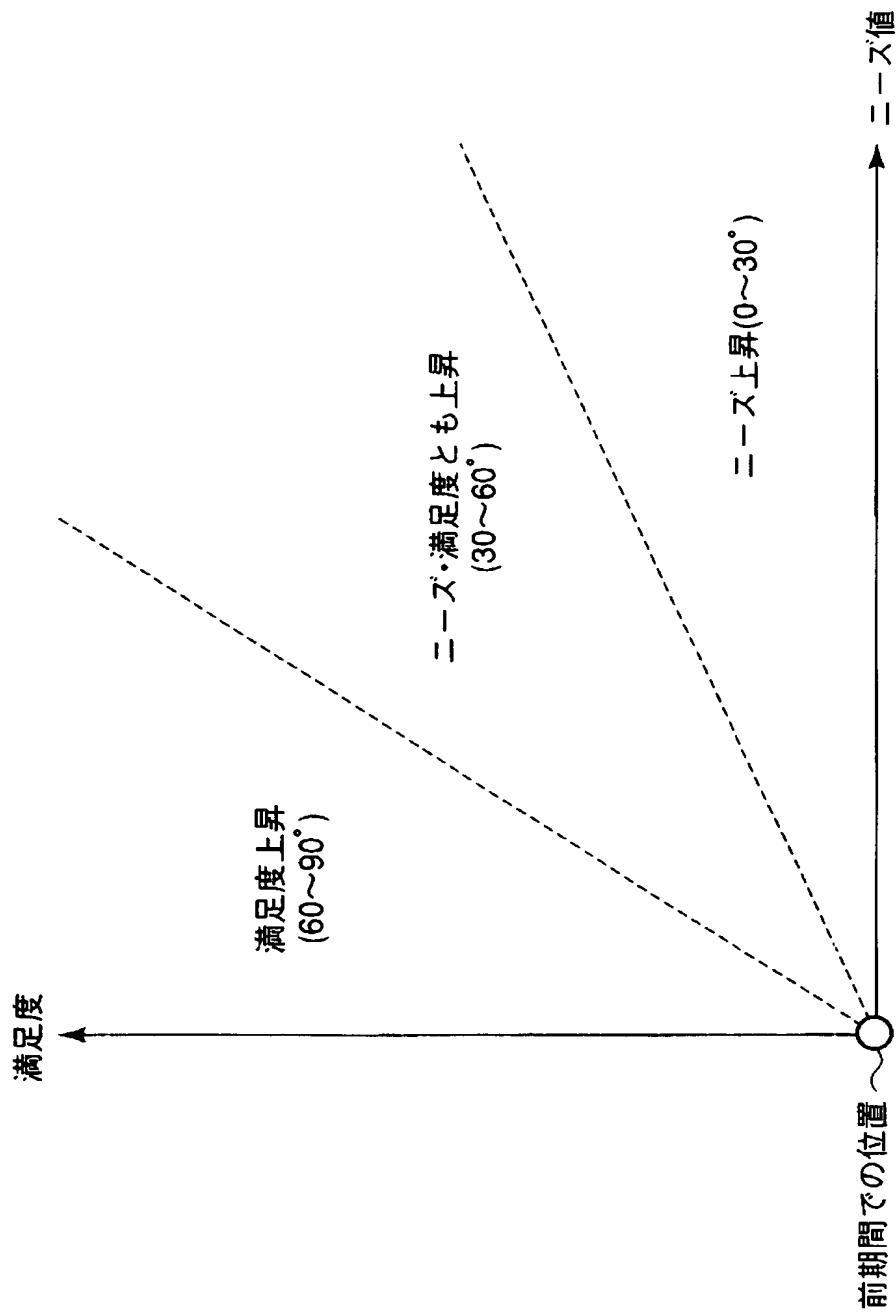
【図26】



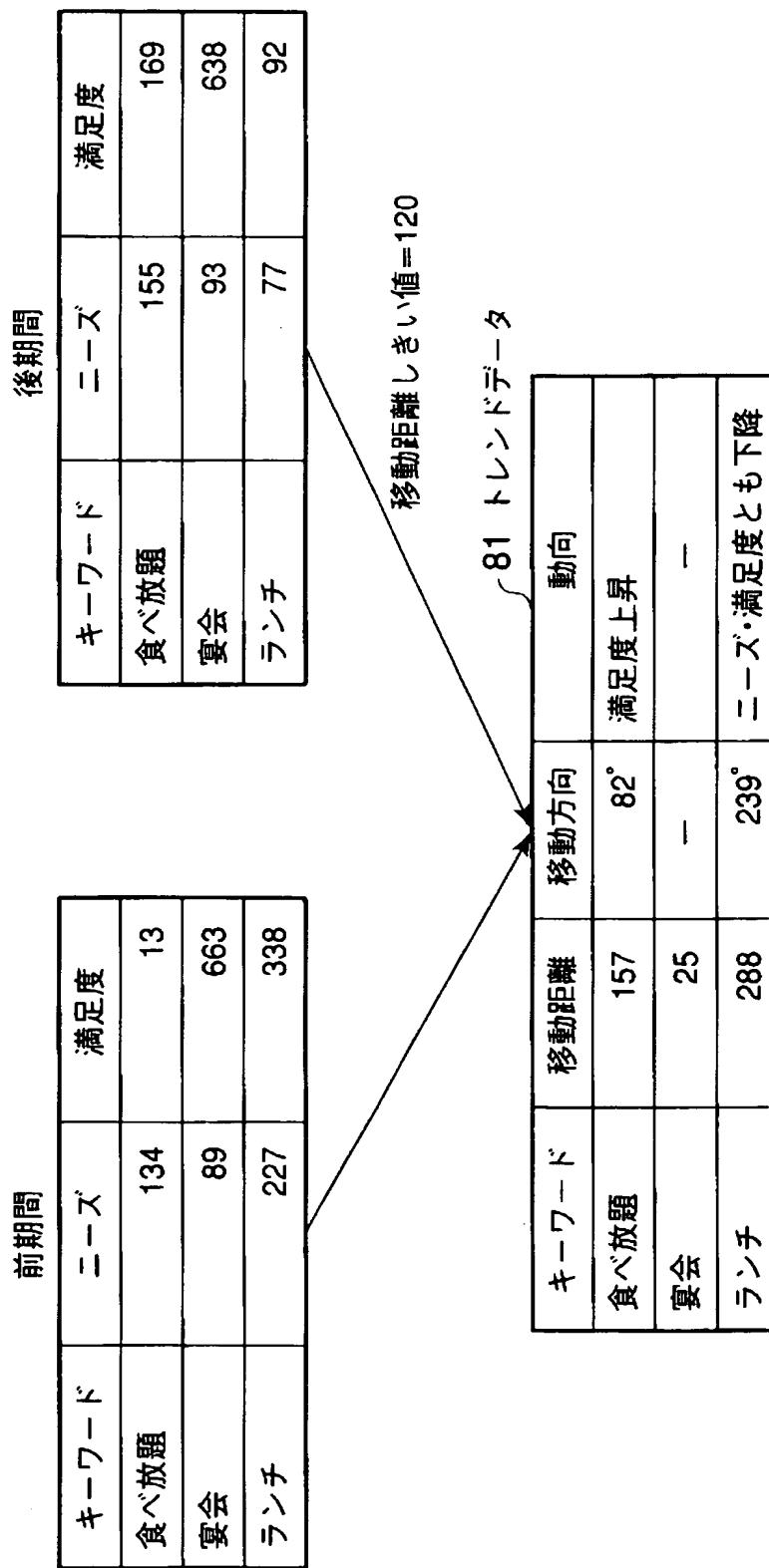
【図27】



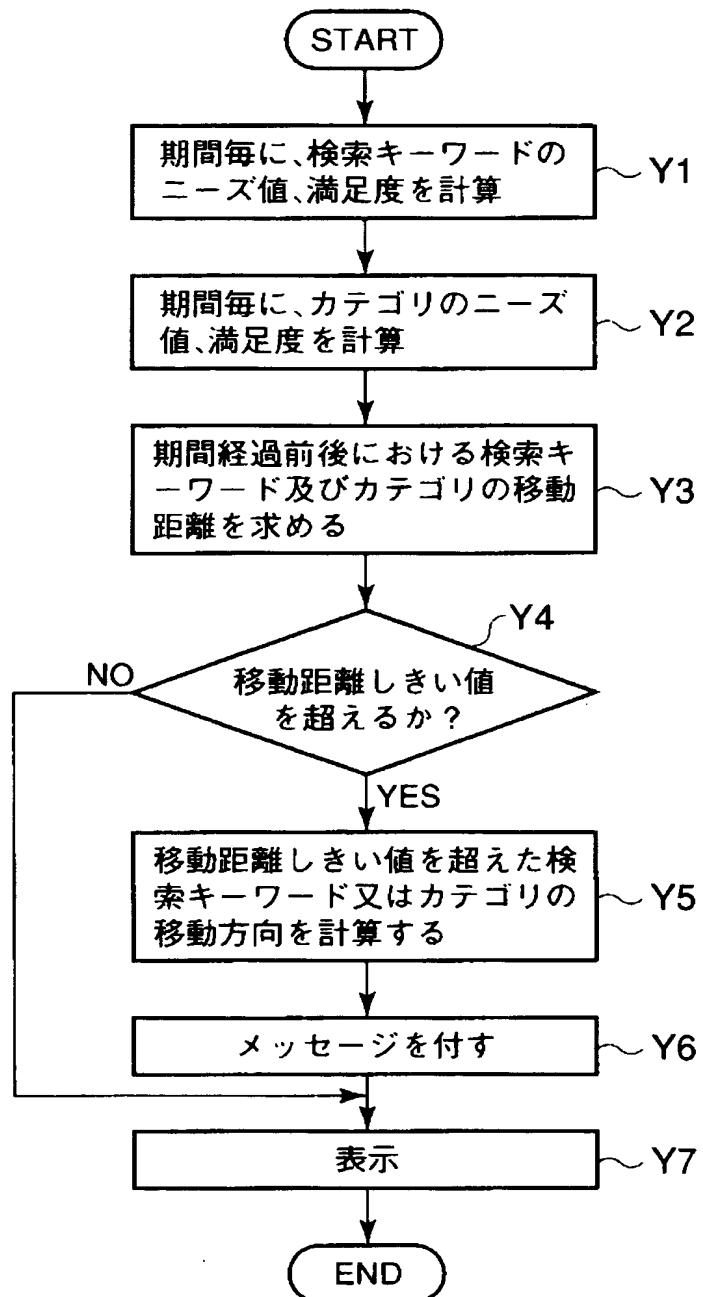
【図 2.8】



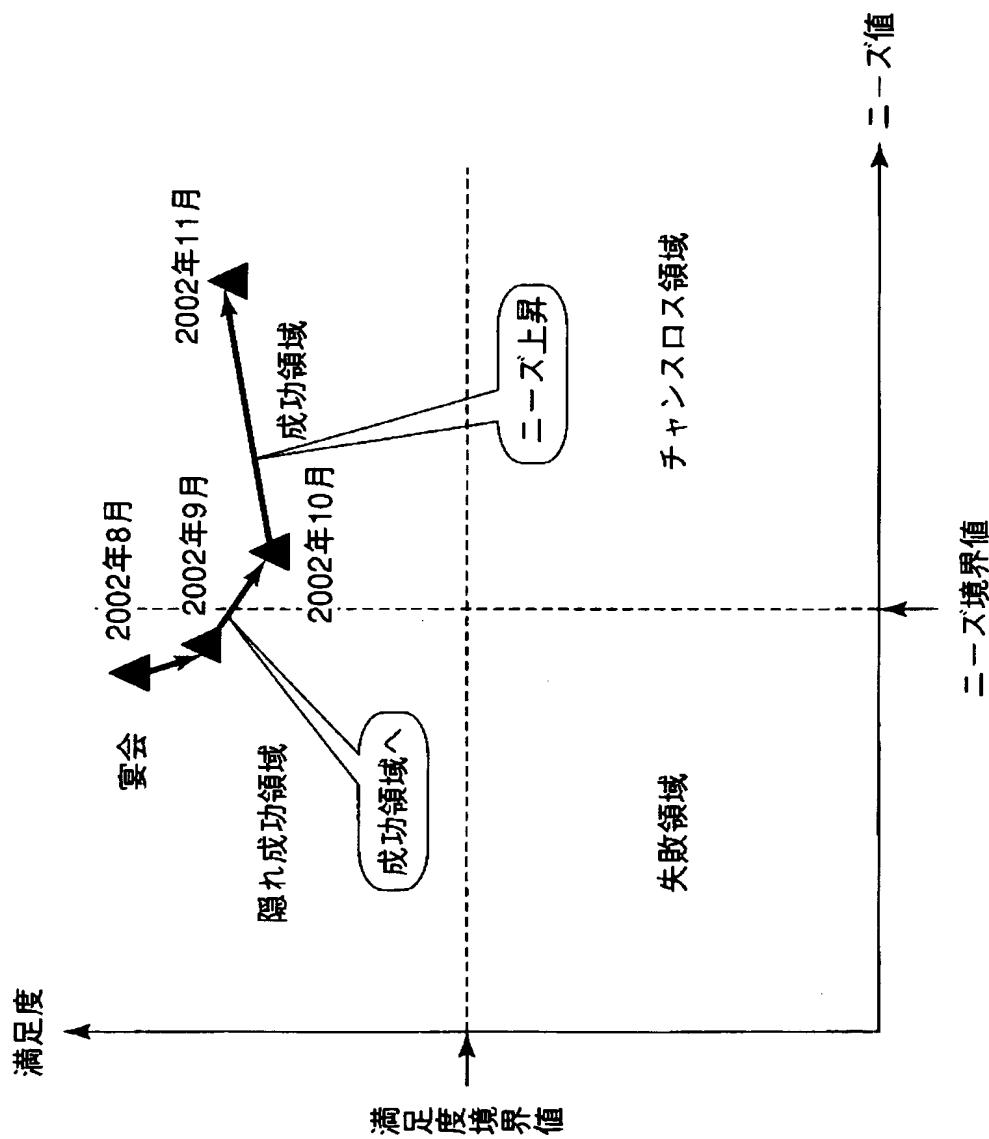
【図29】



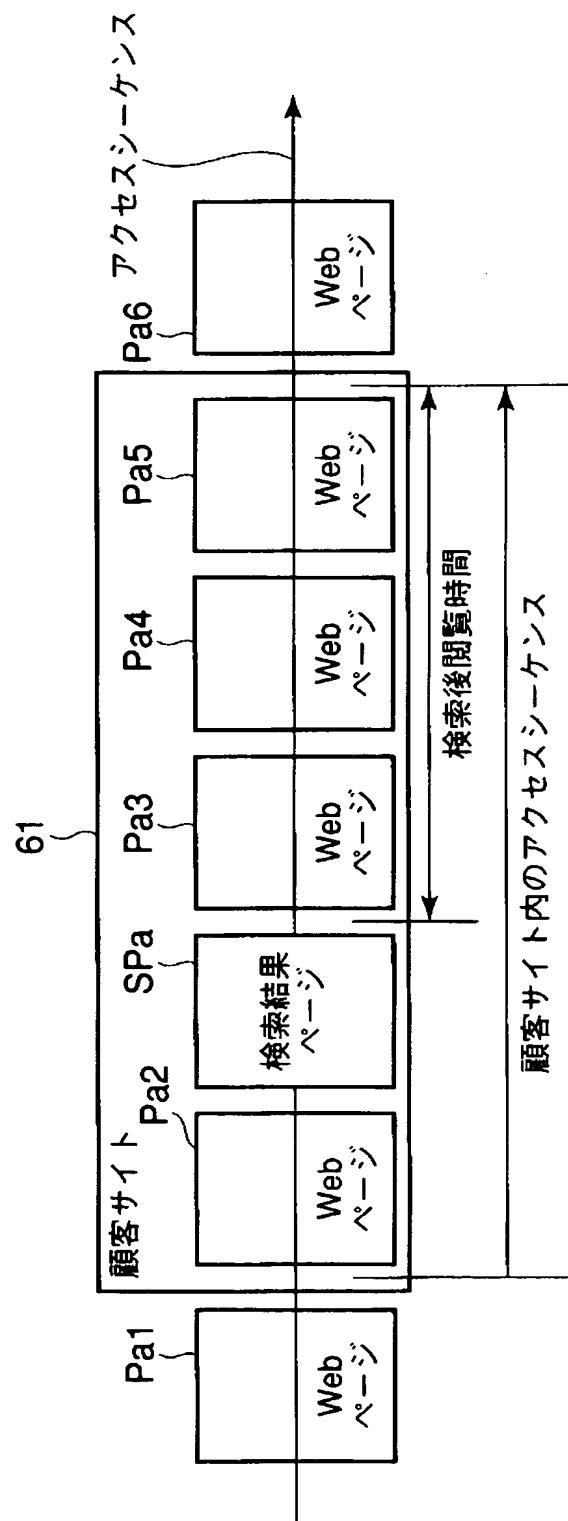
【図30】



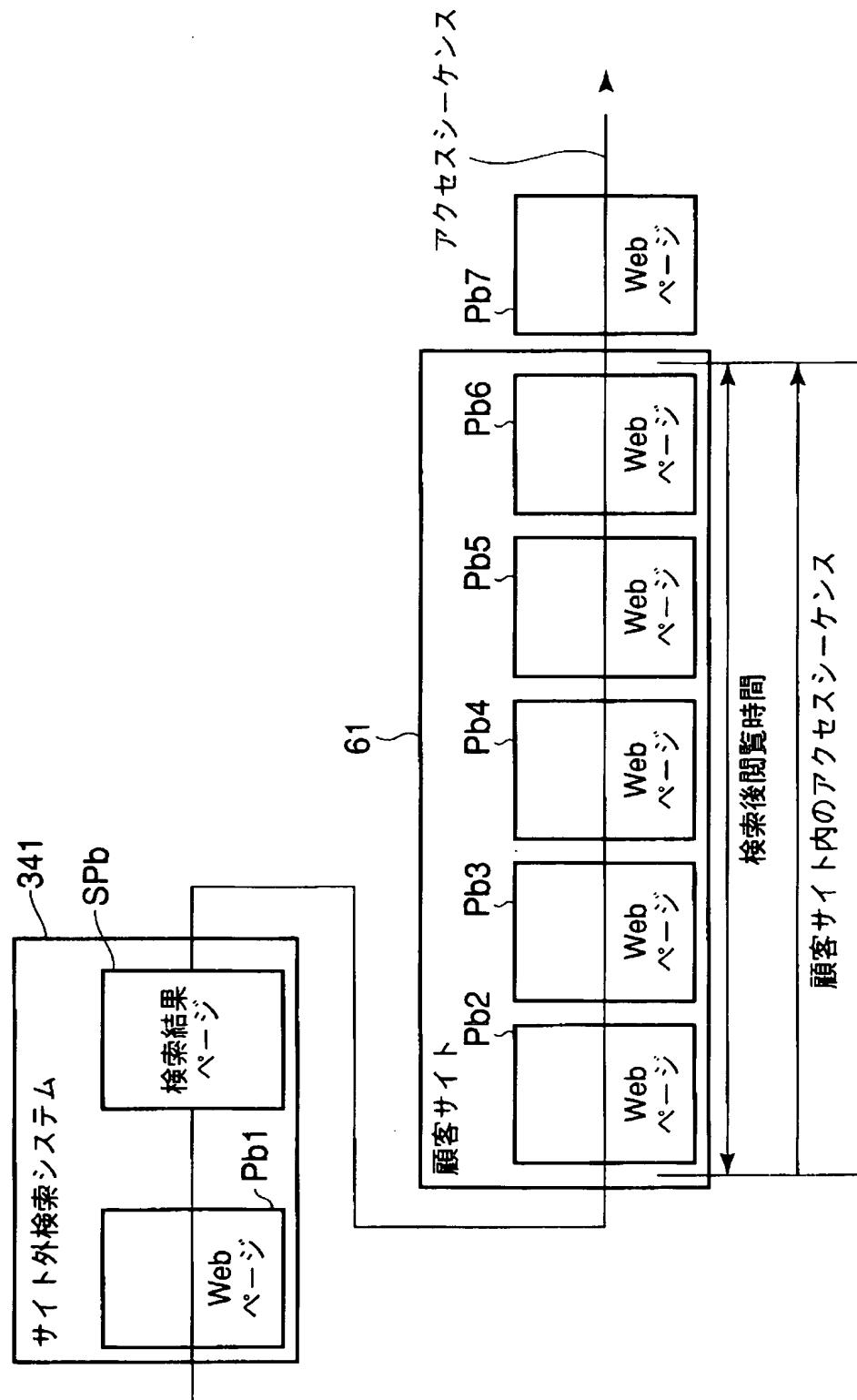
【図31】



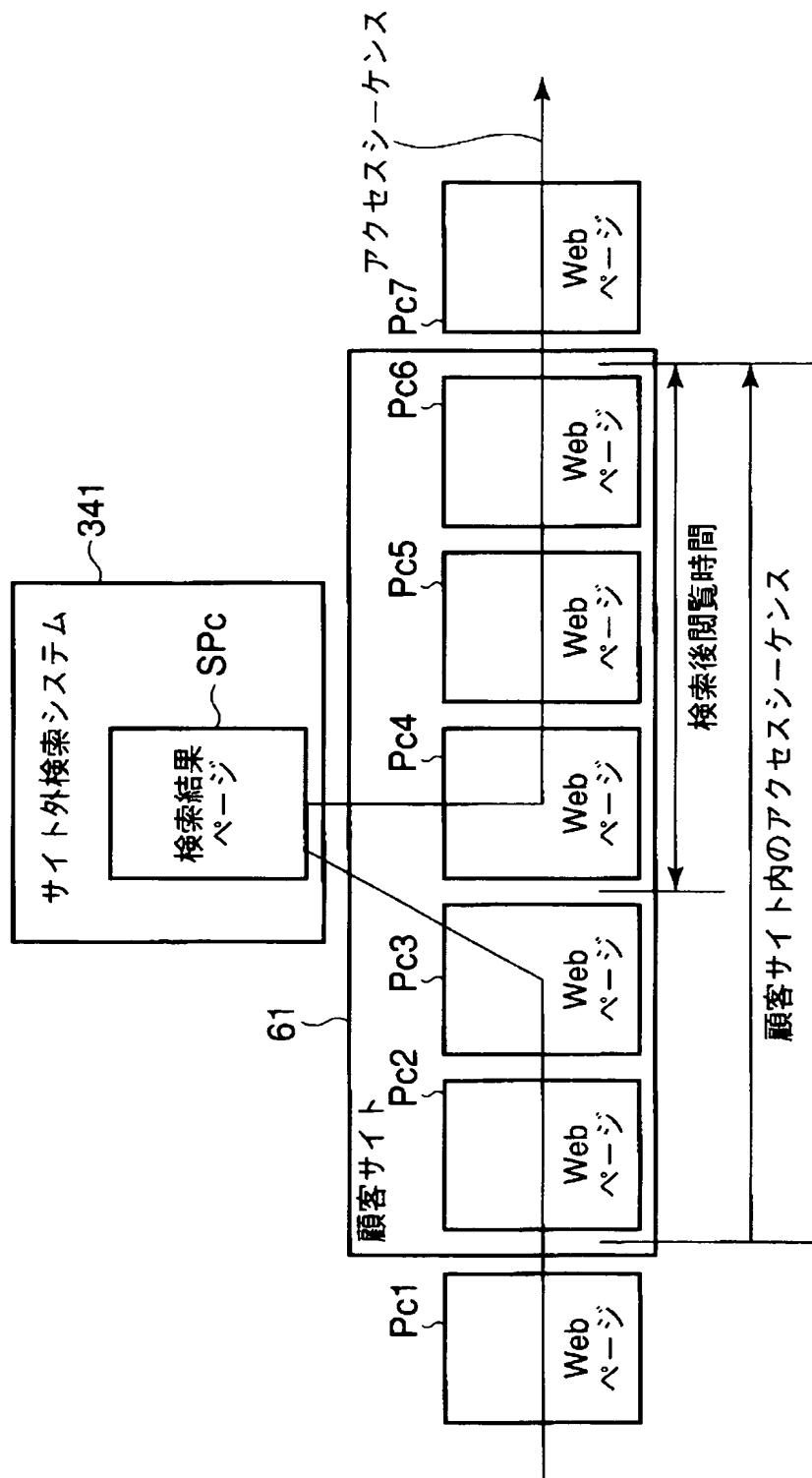
【図32】



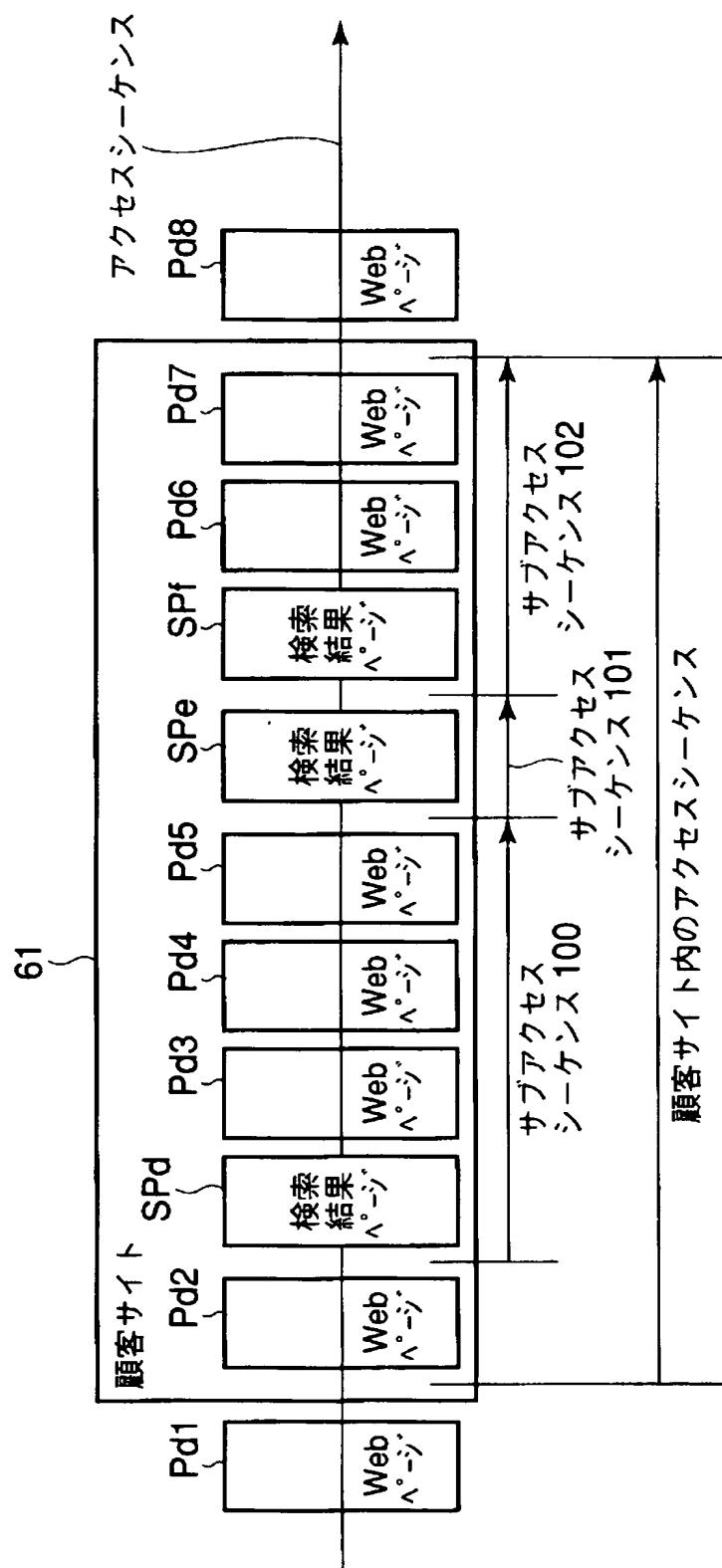
【図3-3】



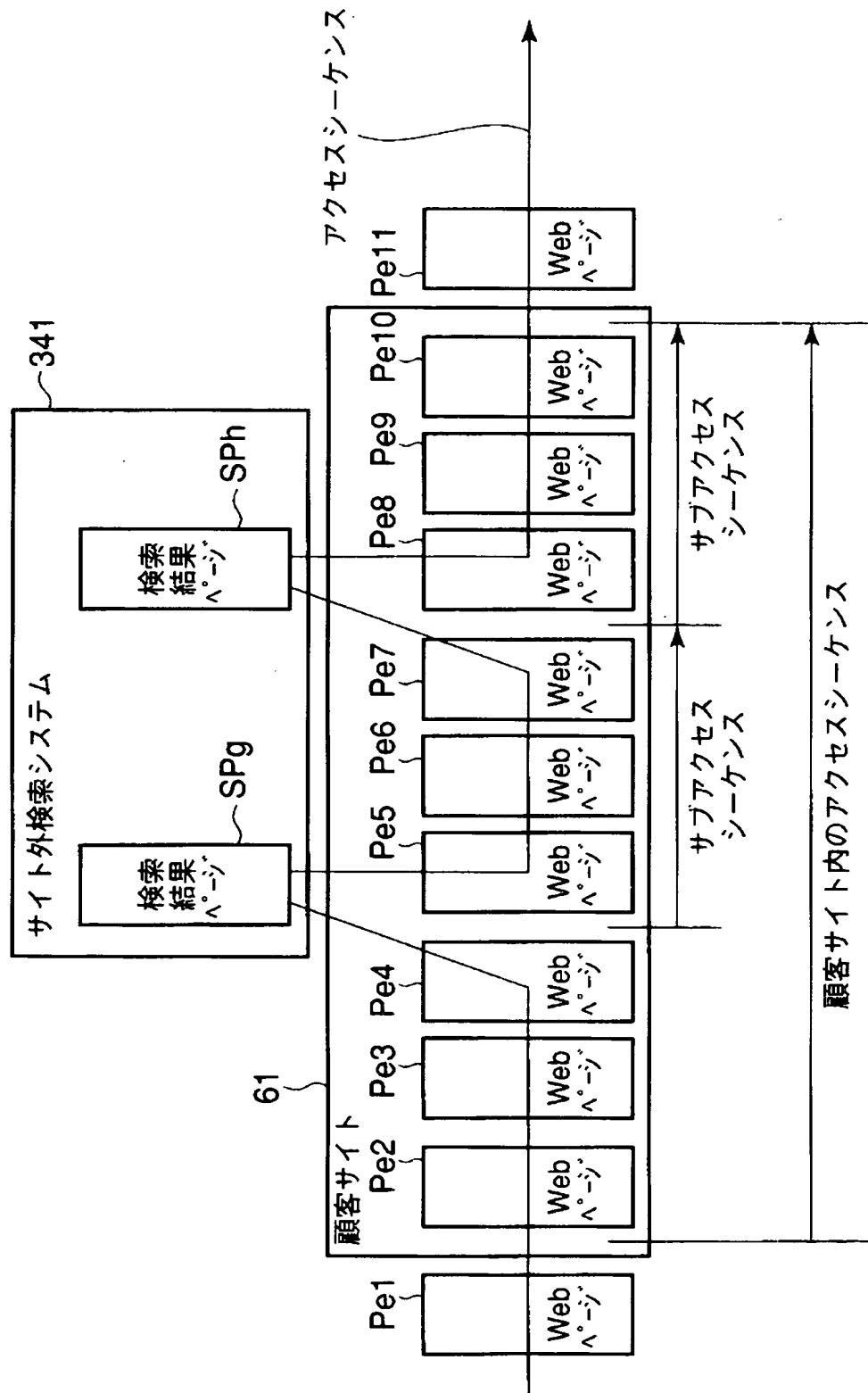
【図34】



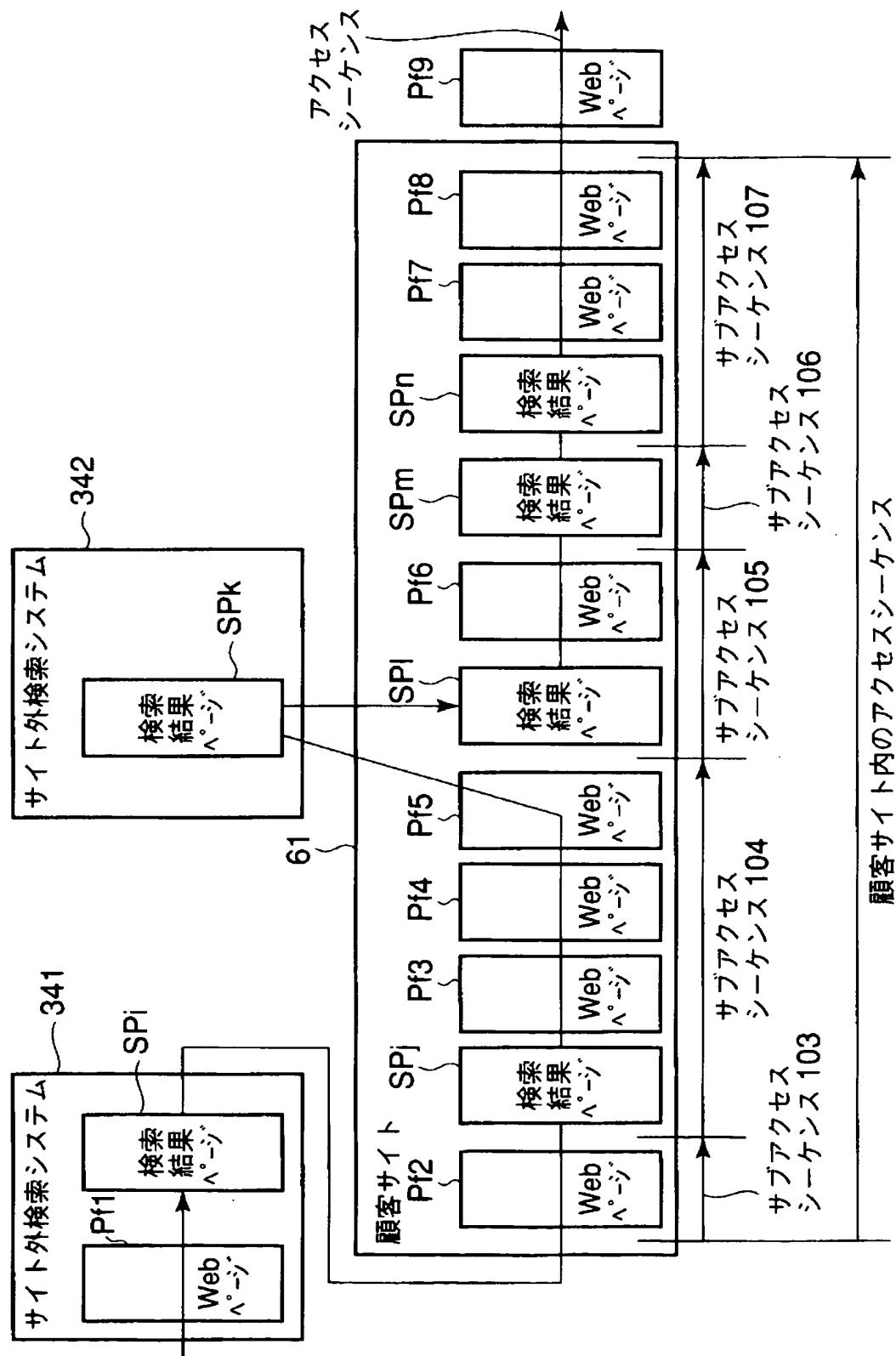
【図35】



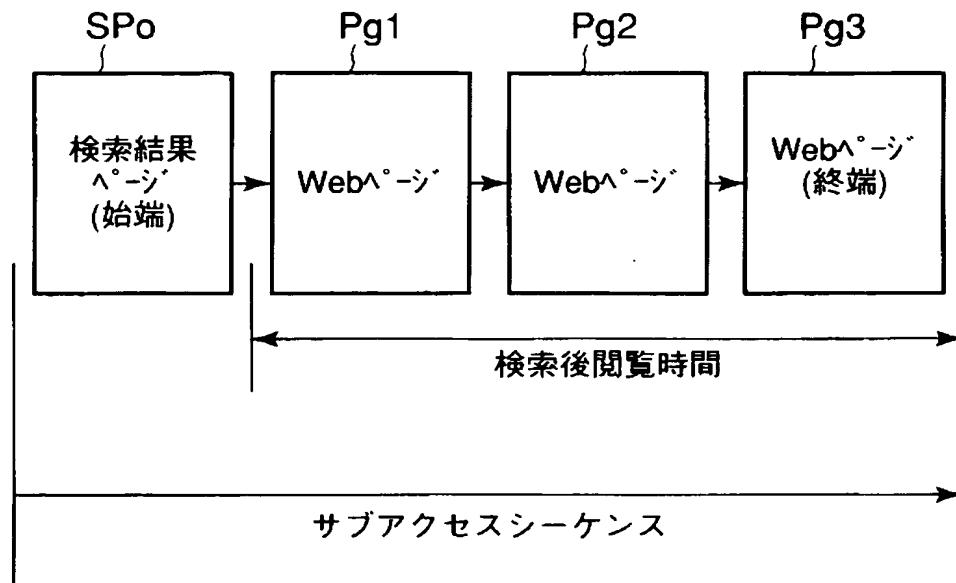
【図36】



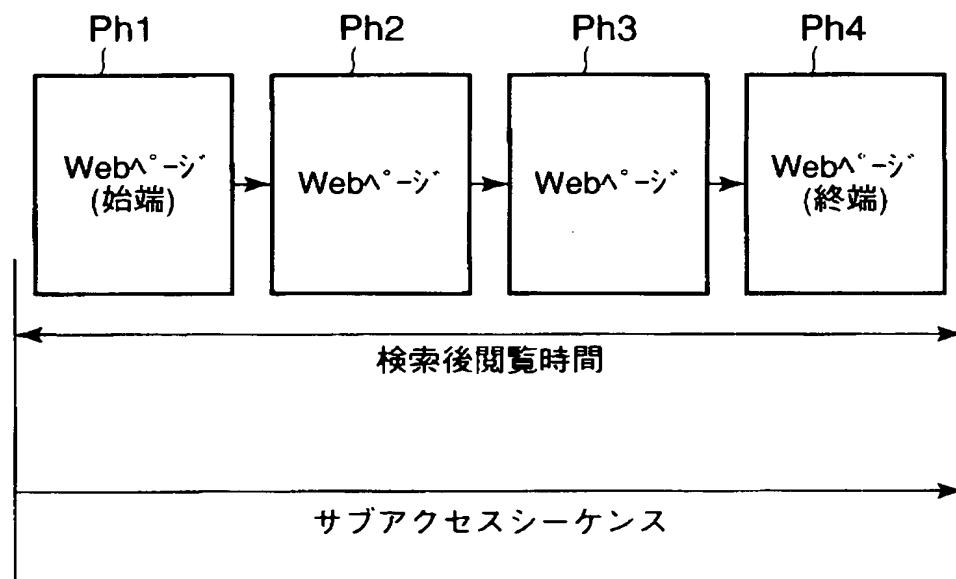
【図37】



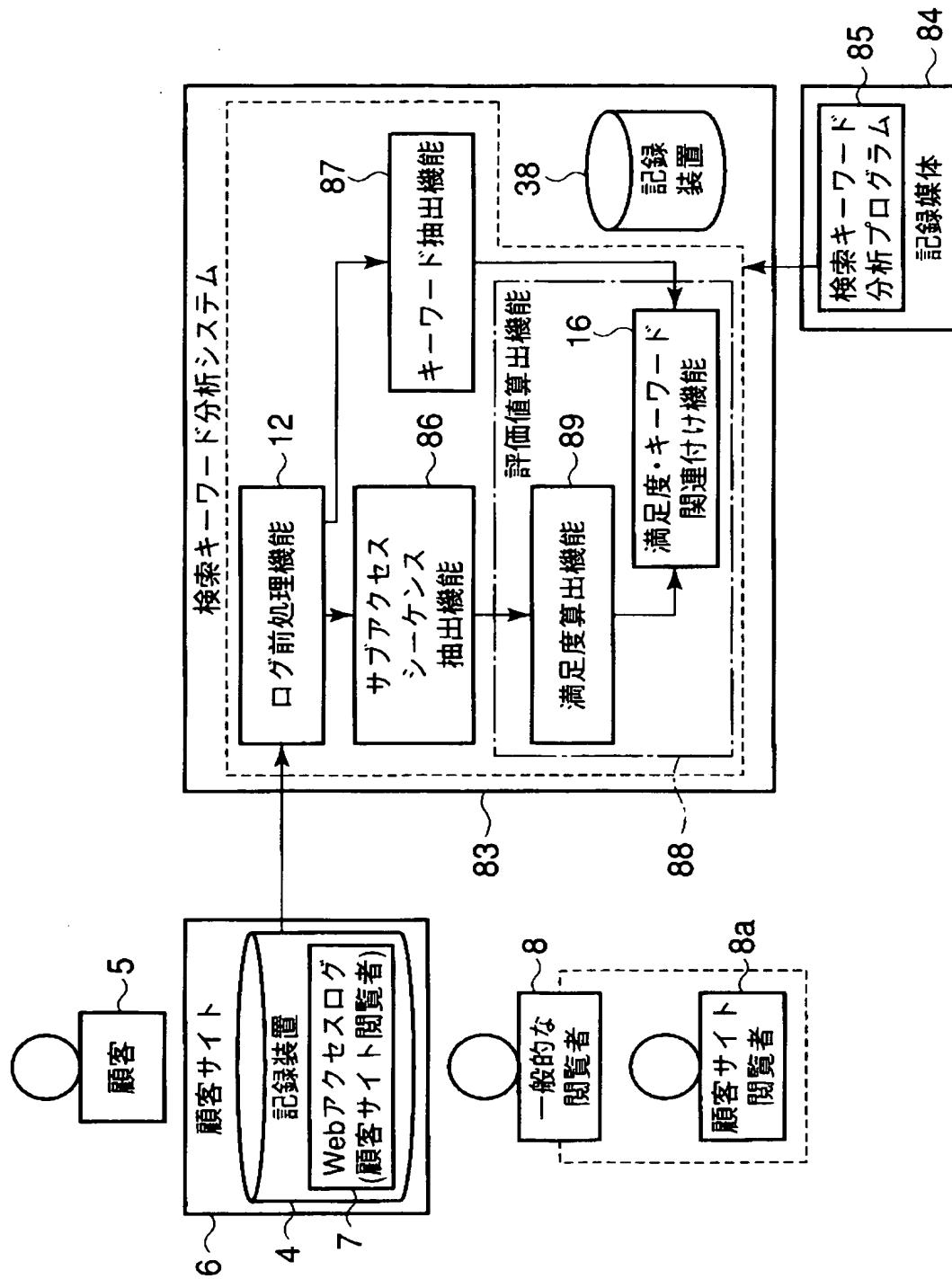
【図38】



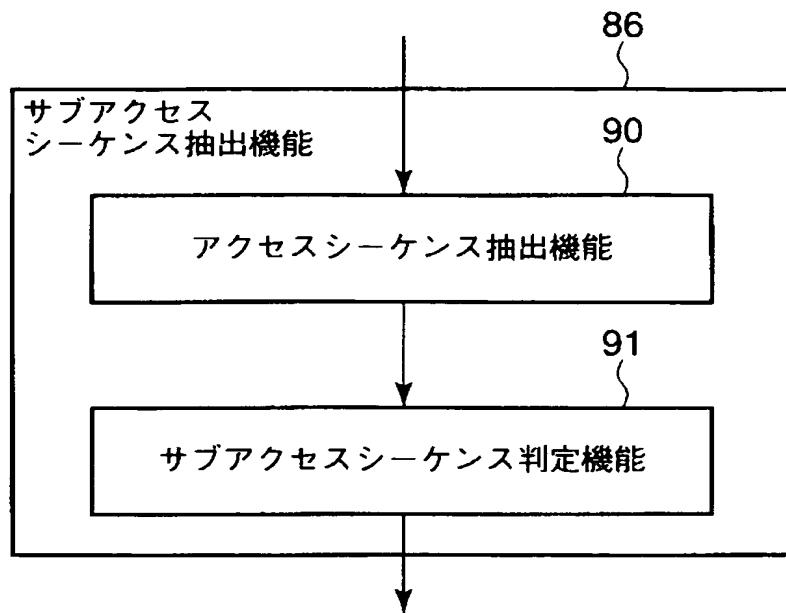
【図39】



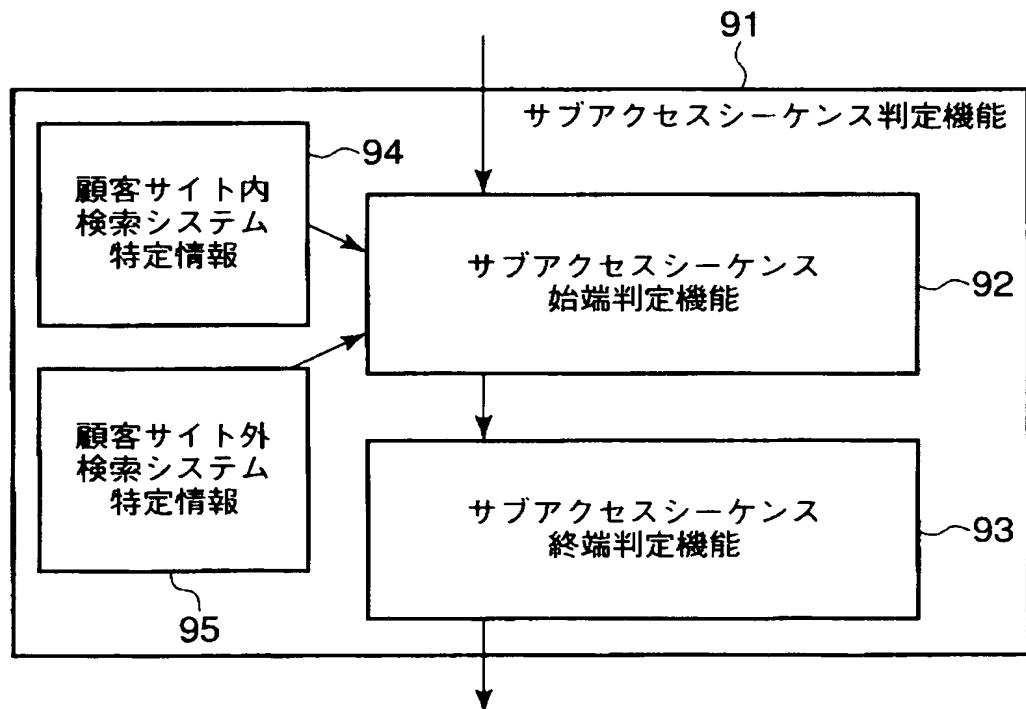
【図40】



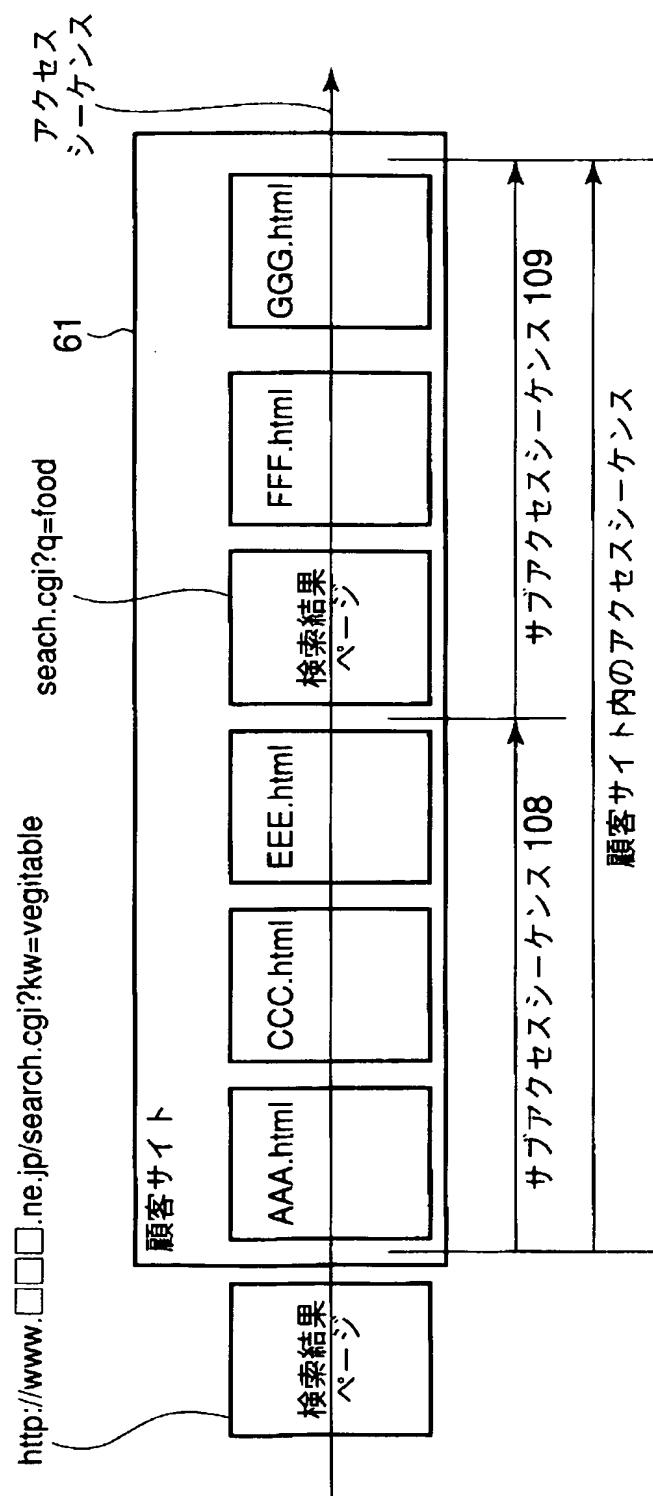
【図 4 1】



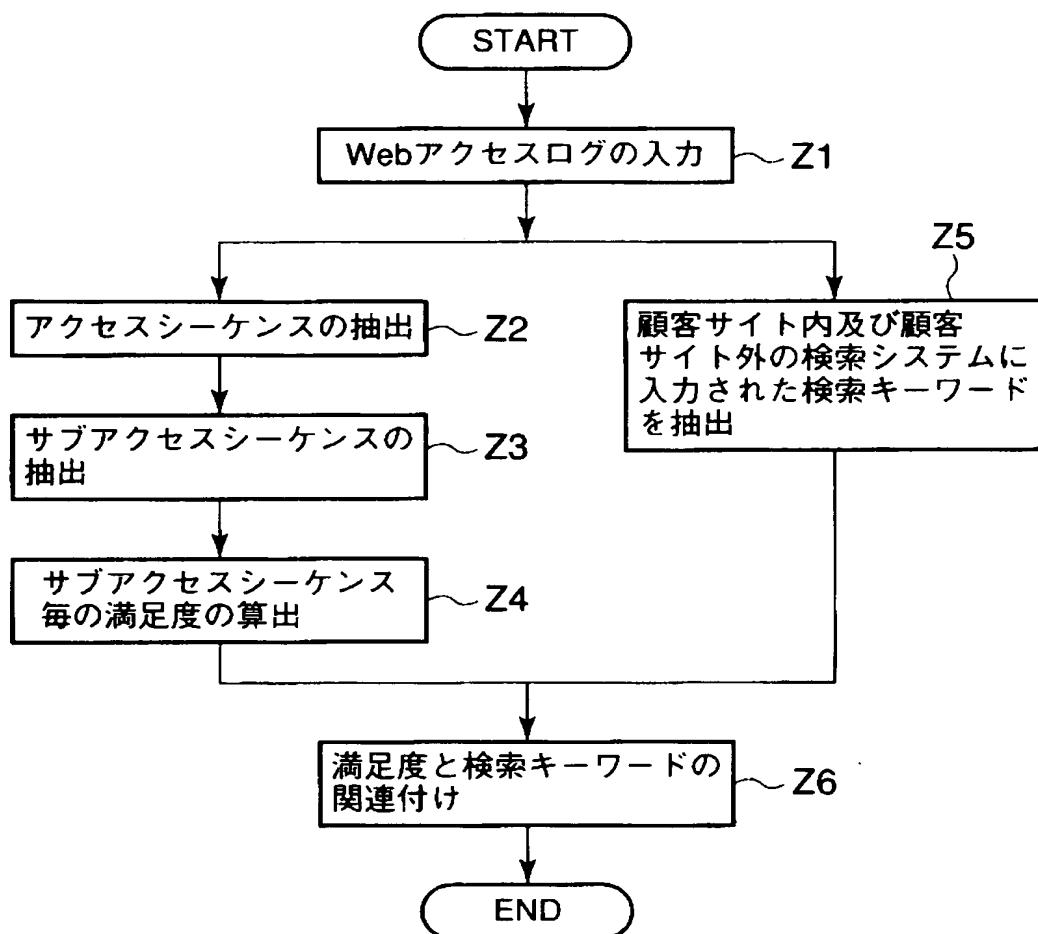
【図 4 2】



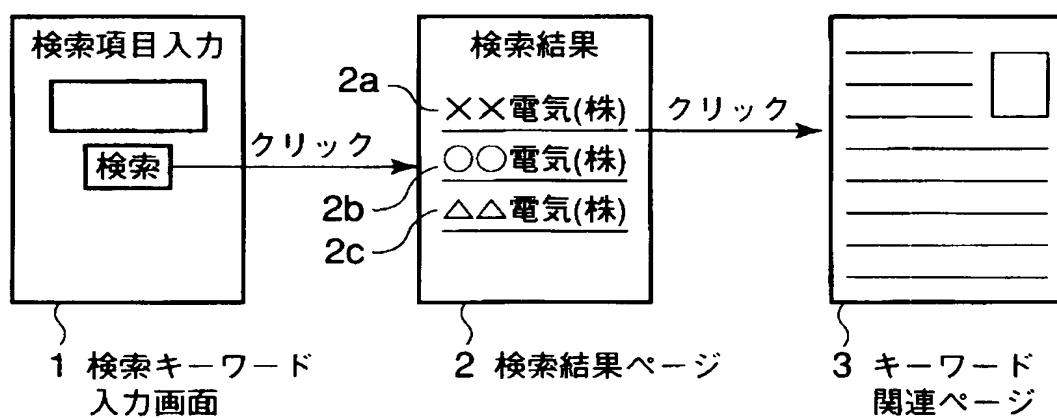
【図 4-3】



【図 4-4】



【図45】



【書類名】 要約書

【要約】

【課題】 入力された検索キーワードに基づく検索システムの検索結果に対する質的な評価値を求める。

【解決手段】 本発明のプログラム11は、コンピュータに、アクセス要求側を示すアクセス要求側識別データとアクセス要求側にアクセスされたアクセス先を示すアクセス先識別データとを関連付けたWebログ7に基づいて同一のアクセス要求側による連続的なアクセスを表すアクセスシーケンスを抽出する機能13と、Webログ7において検索システムを示すアクセス先識別データと入力された検索キーワードとが関連付けられている場合に、検索キーワードを抽出する機能14と、アクセスの状態から検索システムの検索結果に対する質的評価値を算出する基準を定義した質的評価規則に基づいてアクセスシーケンスに対応する質的評価値を算出し、検索キーワードと質的評価値との対応関係を示す関係データを求める機能37とを実現させる。

【選択図】 図1

認定・付加情報

特許出願の番号	特願 2003-070281
受付番号	50300423389
書類名	特許願
担当官	第七担当上席 0096
作成日	平成 15 年 3 月 19 日

<認定情報・付加情報>

【特許出願人】

【識別番号】	000003078
【住所又は居所】	東京都港区芝浦一丁目 1 番 1 号
【氏名又は名称】	株式会社東芝

【代理人】

【識別番号】	100058479
【住所又は居所】	東京都千代田区霞が関 3 丁目 7 番 2 号 鈴榮特許 綜合法律事務所内
【氏名又は名称】	鈴江 武彦

【選任した代理人】

【識別番号】	100091351
【住所又は居所】	東京都千代田区霞が関 3 丁目 7 番 2 号 鈴榮特許 綜合法律事務所内
【氏名又は名称】	河野 哲

【選任した代理人】

【識別番号】	100088683
【住所又は居所】	東京都千代田区霞が関 3 丁目 7 番 2 号 鈴榮特許 綜合法律事務所内
【氏名又は名称】	中村 誠

【選任した代理人】

【識別番号】	100108855
【住所又は居所】	東京都千代田区霞が関 3 丁目 7 番 2 号 鈴榮特許 綜合法律事務所内
【氏名又は名称】	藏田 昌俊

【選任した代理人】

【識別番号】	100084618
【住所又は居所】	東京都千代田区霞が関 3 丁目 7 番 2 号 鈴榮特許 綜合法律事務所内

次頁有

認定・付力口青幸及（続巻）

【氏名又は名称】 村松 貞男
【選任した代理人】
【識別番号】 100092196
【住所又は居所】 東京都千代田区霞が関 3 丁目 7 番 2 号 鈴榮特許
綜合法律事務所内
【氏名又は名称】 橋本 良郎

次頁無

特願 2003-070281

出願人履歴情報

識別番号 [000003078]

1. 変更年月日 1990年 8月22日
[変更理由] 新規登録
住 所 神奈川県川崎市幸区堀川町72番地
氏 名 株式会社東芝

2. 変更年月日 2001年 7月 2日
[変更理由] 住所変更
住 所 東京都港区芝浦一丁目1番1号
氏 名 株式会社東芝